Vol. 32, No. 2 February, 1935

THE
PSYCHOLOGICAL BULLETIN

THE PRESENT STATUS OF THE PROBLEM OF INDIVIDUAL DIFFERENCES IN ALTERNATING ACTIVITIES

BY CLARA BURRI
University of Chicago

Introduction

The psychological concept of individual differences in ability to alternate from task to task is frequently expressed in the popular idea that some people can easily engage in varied activities while others have essentially "one-track minds". Psychologists have concerned themselves seriously with this problem of individual differences since the early part of the 20th century. Investigations relating to it fall into several groups. There are the studies of the rôle of phasic or postural reactions which furnish the background for the particular movements which a person makes from moment to moment; studies of shift of attention; the effect of previous activity on subsequent activity; studies on sensory adaptation and persistence of after-sensation; the effect of repetitive work versus varying work; studies regarding quantitative versus qualitative set; studies of perseveration or set as a general factor; and, finally, studies on individual differences in ability to alternate activities, which, according to some authorities, is determined by the degree of perseveration, or the tendency of activities to persist.

The phenomenon of "perseveration of an activity" operates in the case of the student in the classroom when he must shift quickly from one subject to another; the football player who hears the expression "fifteen yards", a term which will have an entirely different meaning, depending on whether he is on the football field or in a geometry class; and the person, who, when suddenly spoken to in a foreign language with which he may be ever so familiar, is at first

unable to answer or even to understand what has been said to him, because he expected to hear a different language. Again, the influence of perseveration may become a problem in the office or the factory, or, in fact, in any place where the worker is required to change quickly from one activity to another. Sometimes it even appears in the form of an "idea" or "tune" which recurs persistently "in one's mind" to such an extent that all our efforts to forget it seem to be without results. The tendency seen in these illustrations becomes especially strong in certain cases of insanity. In fact, students of psychopathology were the first to call this abnormal phenomenon by the name of "perseveration". Later, psychologists observed that the tendency of activities to persist appeared also, though less strongly, in normal subjects. In 1898 Ziehen (24) applied the term "perseveration" to the disposition which he observed in normal boys from eight to fourteen years of age to use the same reaction word repeatedly in association experiments. Even before that, in 1875, von Kries (11) argued that there are certain cortical functions, "sets", so to speak, which are after-effects of experiences and which will either help or hinder further experiences.

In 1900 Müller and Pilzecker (15) formulated perseveration into a so-called fundamental law of memory. Every "idea", they said, has its perseverative tendency which diminishes quickly with time and returns to consciousness of its own accord. This hypothesis they demonstrated in their work on memory, in which they showed that syllables which had been given as correct responses would in the same series later appear as incorrect responses. Some syllables would appear as responses for several days. This perseveration of syllables was found to be stronger with fresh series (ones tested shortly after learning) than with older series. Marked individual differences were noted, some individuals showing much of this perseveration of syllables, and others having little or none of it. From this discovery there arose two groups of independent studies, one of which aimed at determining some of the factors which influence the hindering or helpful effect of one activity upon another and at finding the retroactive effects of activities. The fact that some scholars doubted the correctness of Müller's finding that an activity tends to persist over several days formed the main reasons for this development. However, we are not concerned with these studies in this paper.

Another group of psychologists undertook to test Müller and Pilzecker's hypothesis regarding individual differences in persevera-

tive tendency. In these studies it has frequently been assumed that the degree of a person's perseverative tendency determines his ability to change from task to task. In fact, the difference between the amount of work performed when tasks were alternated and the amount done during continuous work on one task has often been used as a measure of perseveration. According to this point of view, those people who show little perseverative activity are people who can readily alternate between tasks, while those who have strong perseverative activity change "set" only with difficulty and hence are not as good at alternating work as they are at continuous work.

If we note the widely different approaches to the problem of the influence of alternating versus constant activity, and if we carefully analyze the varied and at times conflicting results, we become aware that ability to alternate between tasks may be explained in various ways. Although it may be true that "non-" or "weak-perseverators" can shift tasks more easily than "strong perseverators" and that "strong perseverators" may even work better at continuous than at alternating activity, it does not necessarily follow that all, or even most, of the differences in the results of alternation versus constant activities must be explained in terms of perseveration. In fact, as we shall see later, a number of group studies of alternation indicate several different factors which determine the ability to alternate tasks.

Thus the problem under investigation here is, in the main, one of individual differences in ability to alternate tasks, a problem which is much more complex than is ordinarily assumed, and which in order to be understood fully, must be viewed from a broader perspective than has hitherto been done. The studies of individual differences form only a small part of the total group of studies which deal with the effect of alternating versus constant activity. It is therefore our aim to view the problem within its setting, and to show that its background furnishes much valuable material for the understanding of individual differences in alternating tasks.

A review of the literature reveals the problem of alternating activities versus constant activity to have been investigated from several different angles. The studies may be classified into two groups according to their main emphasis, the one group examining the effect of alternation, and the other investigating individual differences in ability to alternate. In the past these two groups of studies have not only approached the problem from different angles, but they have treated it as if the two questions were unrelated. As pointed

out in the introductory pages, individual differences in alternating tasks have hitherto been accounted for almost entirely in terms of perseveration. Group studies on the effect of alternation make it clear, however, that various other explanations also apply.

We shall show here that the apparently independent problems are related and that their conflicting results are the outcome of differences in emphasis. In order fully to understand the problem of individual differences in ability to alternate, the findings of the studies from both groups must be combined. In the next section we shall analyze some of those group studies which are most closely related to the problem of individual differences.

The Effect of Alternation

The studies in this group deal in the main with the general or group effects of alternation. Some of the studies were performed in actual work situations, while others consisted of laboratory experiments. In both series of investigations it was found that superior average performance was obtained under alternating conditions. The investigators explained these results, usually, in terms of well-known concepts of general work decrement, such as fatigue, interference of habit patterns, refractoriness, monotony and boredom, work motivation and attitude, and, in a case or two, by the hypothesis of perseveration. First we shall look at certain experiments most closely related to the problem of individual differences in alternating tasks, in order to discover what factors determine the ability to alternate. Later on we shall see that those factors which affected the general or group results of alternating versus constant activity are also influential in causing individual differences.

The Laboratory Studies. Among the first to study the effect of alternation was Weygandt (21). He compared alternating and constant work by six subjects in various activities such as learning digits and syllables, performing additions, and reading Hungarian or Latin prose. Each subject worked at a given task either constantly for one hour and fifteen minutes or for one hour, after which he changed to another task for fifteen minutes and then changed back again to the original task, to work on it for a final fifteen minutes. In order to determine the effect of alternating activity, the author compared the amount of work done on the experimental task during the fifteen minutes immediately following the interpolated activity, with the last fifteen minutes from the one hour and a quarter of continuous work. Weygandt found that alternation at times increases but at

other times decreases efficiency. According to his interpretation the difficulty of the tasks alternated seems to be the major factor in determining whether alternation produces an increase or a decrease in efficiency. When a given task is followed by a more difficult one, there will be a decreased efficiency in the easier task; while if a given task is followed by a less difficult task, alternation produces an increase in efficiency in doing the first or difficult task. Whether a given task is considered easy or difficult seems to depend upon the individual's intelligence and his degree of practice.

We may interpret these results by assuming that during the process of work people develop a certain set or attitude toward the work. This set may be either facilitating or inhibiting, and will influence a person's efficiency in the job accordingly. If this be the case, then an easy activity following a difficult activity should produce self-confidence or similar energizing attitudes, and thus increase activity; while the opposite procedure, that is, an easy activity followed by a difficult activity, should decrease it.

Chapman (4), as may be seen in the report of his work on the effect of rapid change of set on the speed of work, also thinks that the difficulty of the tasks alternated may have something to do with determining the facilitating or inhibiting effect of alternation. However, he points out particularly that the effect of alternation will depend upon whether a mental or a motor activity is employed. With the aid of simple addition and letter cancellation he measured the rate of work during a three-minute period when the subjects either worked continuously at adding or at cancelling letters, or alternated every three seconds between the two activities. Efficiency of work was measured in terms of the amount correctly done during the three-minute period. The subjects were 14 per cent more efficient in adding when they were alternating it with cancelling letters than when they were adding continuously. When they changed from cancellation to addition, however, the effect of change was negligible. This difference in results is explained in terms of interference. In the case of continuous cancellation, an activity which seems to be chiefly motor, the previous cancellations do not have to be given time to fade away to any great degree. When the subject changes from problem to problem of addition, which is chiefly a "mental function", he must forget all previous results except the one necessary for the immediate calculation. During cancellation the old figures have time to fade away, and therefore more work can be done when alternating. The difference between the two kinds of tasks may also be due to the

degree of difficulty. Cancellation, for example, may become automatic. Although Chapman seems to favor an explanation in terms of the type of the task, whether mental or motor, he agrees that the difference may perhaps be due to the fact that the mental function was more difficult than the motor function. Chapman's results can easily be explained in terms of perseveration, simply by substituting the "perseveration hypothesis" for interference. Perseveration in this case would apply to the tendency for specific numbers to recur.

Dashiell (6) used different conditions of alternation than did Chapman, and consequently got different results. He used 100 simple arithmetic problems, twenty-five each in addition, subtraction, multiplication, and division, which he presented to 69 subjects either in mixed order, or continuously, first all of one kind and then all of another. From this group of subjects "63 took less time for completing the continuous than for completing the mixed order of examples; the group averages being, respectively, 159.3 seconds and 181.7 seconds. The recall of a specific habitual addition is more prompt when one has just been adding than when he has just been doing a variety of things. So with subtracting, with multiplying, with dividing" (6, p. 282). This the author attributes to set or persistency of activity. This explanation may also be referred to as a perseveration effect, but in this instance it is a tendency for the "work set" to persist or continually to reassert itself.

Jersild (8) reports some experimental results obtained from the study of the effect of different kinds of tasks on the ability to alternate. He tried to see whether persistency of set and ability to shift are independent factors. He gave ten different tests to groups of from twenty to thirty subjects, in which they had to shift from one task to another, or to work continuously. The ten different tasks gave widely varied results. For some combinations of tasks, working alternately decreased efficiency, while for other combinations, it increased efficiency. These findings seem to warrant the general conclusion that the effect of alternation varies according to the type of task alternated.

As a supplementary part of investigations in actual work situations Wyatt (22) also performed some laboratory studies. In these experiments three subjects either worked continuously for two and one-half hours in the morning and afternoon, or alternated every fifty minutes. Three kinds of activity were used: (1) simple addition, (2) mechanical computation, in which columns of digits were added by means of a comptometer and the results recorded on a typewriter, and (3) muscular work, in which the subjects pulled against

a spring balance with the left and the right hand alternately. The increase in total output under alternating conditions varied from 2 to 24.2 per cent; the number of errors decreased from 9.2 to 5.1 per cent. It also appears that for each form of activity an optimal frequency of change exists, which probably depends on the type of work and the individual peculiarity of the worker. The results indicate that the same objective conditions may have different subjective effects on different persons. There are differences in ability to alternate and also differences in individual liking or disliking of the alternating and the constant method. It is possible that the individuals who find repetitive work most monotonous and who do better and more consistent work by alternation possess organic and nervous peculiarities which differ from those subjects who work better continuously.

Robinson and Bills (17) had 24 subjects work both continually and alternately on simple problems such as writing or reading letters and tapping with the fingers. In general their results, which are group totals only, agree with those obtained by other investigators, namely, that on the average alternating work produced greater efficiency than continuous work. However, in the finger movement experiment in which the subjects tapped with the first, second, or third finger or successively with two or three of these fingers upon the keys of a typewriter, it was found that alternation was no longer advantageous to performance. It seems that with nicely balanced behavior systems, continuous activity may produce greater efficiency than alternating activity.

These authors' work is particularly worth considering because of their explanation of the superiority of alternating activity. Robinson and Bills apply the physiological concept of refractory phase to the psychological phenomena which they observed. Just how this refractory phase phenomenon should work in connection with alternating and constant activities may be best seen from the authors' own words. They say:

> The large biological phenomenon of refractoriness is a direct correlate of the recency and frequency of functioning of the activity in which the refractoriness is manifest. Since relatively homogeneous work involves greater recency and frequency of functioning of its part activities than does relatively heterogeneous work, we should expect psychological activities of a homogeneous sort to manifest more refractoriness and therefore more decrement than work of a more heterogeneous sort. (17, p. 422.)

At this point the reader must be cautioned against interpreting refractory phase in the usual physiological way, as it was used by

Adrian, who employed the term to describe the latent period following stimulation of a nerve fiber or nerve ending in muscles and sense organs. Robinson and Bills use the term "refractory phase" in a much broader and looser sense, in the fashion of Verworn and Dodge.

A study made by Bills and McTeer (2) in which they investigated the effect of identical elements on transfer of fatigue, suggests in an indirect way that the similarity of tasks alternated influences the degree of effect from alternation. They found that decrement from alternating two tasks as compared with that from continuous work is proportional to the number of identical elements in the two tasks. This indicates that tasks which are very different would give greater increase in efficiency when alternated than would very similar tasks.

Before leaving the discussion of laboratory investigations and turning to a survey of studies carried on outside the laboratory in actual work situations, let us notice briefly the results of the effect of alternating more than two tasks. This problem has only been touched upon in past studies. Both Jersild (8) and Zillig (25) introduced, along with their other conditions, a few such sittings. For example, the subjects drew triangles, then performed additions, then associations, etc.; then again triangles, additions, etc. In Jersild's study, alternation took place after every element; in Zillig's, after every three to five elements. Jersild reports a gain for alternation in nine out of twelve trials. However, he does not tell the amount of gain. Zillig reports that some children do better when alternating and others when working continuously.

To summarize the findings from the laboratory experiments on ability to perform alternating versus constant tasks we may say that: (1) the general direction of the results is toward superior average performance for alternating activity, though in some situations people seem to work better continuously than alternately; (2) according to the studies which we have just discussed, ability to alternate seems to be influenced by several factors, among which may be mentioned the kind of task, the attitude of the subject toward the task, and the persistence of set; (3) superiority in efficiency during alternating work has been explained thus far in two ways, namely, as being due to less refractory phase effect during alternation than during constant repetition of the same task, or as resulting from absence of interference due to persistency of specific responses which lower a subject's efficiency during constant work; (4) inferior average performance for alternating as against continuous work was explained in terms of "perseveration", according to which hypothesis persistence of a general set interferes with the process of changing from task to task.

Let us turn now to studies which were performed in actual work situations, in order to discover what light they throw on the problem of ability to alternate.

Investigations in Actual Work Situations. Wyatt (22) has performed the most comprehensive studies on the effect of alternation upon activity. He investigated the influence of uniform versus varied work in the packing of drugs, where he used three plans of work. In the first, the drugs were assembled, counted, packed, and wrapped by each worker. In the second, the task was changed every half-hour, and in the third, the same work was done for the whole day. The half-hour changes proved least effective. Plans one and three gave about the same efficiency when measured in terms of actual work performed, but the reports of the workers indicate boredom and monotony for continuous work. Wyatt gives no specific explanation of this, but from the findings of these and of several other investigations he concludes that the interval between alternations tends to influence performance during alternating activity.

In another study, Wyatt and Fraser (23) compared alternating and constant work in soap-wrapping, handkerchief-folding, tobacco-weighing, making cigarettes, assembling of bicycle chains, and assembling cartridge cases. In soap-packing the subjects had either to finish every piece of soap, wrapping, gumming, and labeling it and putting it into a box (alternating activity), or to do first one and then another of these tasks until 144 tablets of soap were packed. Six girls were studied for five days, the results showing that in the case of each worker continuous work was superior to alternating work, the average rate of working under the former condition being 3.3 per cent quicker than under the latter, and varying from .5 to 6.2 per cent for different workers.

In the handkerchief-folding test eight girls were studied for a period of three weeks. Here, two styles of folding were used: the girls followed either one method or the other continuously, or alternated between the two methods after every dozen handkerchiefs. In this task alternation, on the average, increased efficiency slightly. However, if the individual results are analyzed it will be seen that two subjects worked better continuously. For assembling bicycle chains, where four subjects were studied, the results came out in favor of alternation.

Cigarette-packing was divided into making and packing. The subjects either made cigarettes all day or they alternated between making and packing according to their own inclination; or they made them for one hour and then cut; or they made them for one hour

and one-half and then changed to cutting. Each of ten workers was observed for about two weeks in the course of a ten-week period. The number of cigarettes made per five minutes was noted. Again the average output was higher for alternating than for constant work.

The results of these studies indicate that constant work is, on an average, less effective than alternating work. However, the degree of difference in efficiency between alternating and constant work seems to depend upon the type of work and the frequency of change.

In order to study the effect of the frequency of change the authors had four subjects weigh or wrap tobacco all day, or alternate between weighing and wrapping it. The results show that when the work was alternated, the output was higher than when continuous work was performed. This was true for all the subjects. The average increase was 13.8 per cent when two changes were introduced within about four to four and one-half hours of work, and 11.2 per cent when only one change occurred. Individual differences in response to alternation were great. The increase in daily output varied from 5 to 25 per cent. Miles and Skilbeck (**14**) obtained similar results. They found that by introducing a " change period " of fifteen minutes about the middle of the morning and afternoon period of work, during which the girls collected the material required for packing, it was possible to increase the total output by 14.2 per cent.

That the type of work influences the effect of alternation was shown further by Wyatt and Fraser (**23**) in an experiment on the assembling of cartridge cases, which consisted of stacking, assembling, and pushing down the wads. These activities seem to represent different types of work, differing particularly in the degree of fatigue which they produce.

Wyatt and Fraser believe that the increase in efficiency during alternation of work, as compared with constant work, results from the fact that local fatigue effects are relieved by alternation. They maintain that fatigue produced by light repetitive work is more local than general. A suitable change in the form of activity seems to utilize comparatively unused parts of the bodily mechanism, thereby providing a situation which is favorable to productive activity. Hence, they insist, if the most satisfactory results are to be obtained, a muscular process, for instance, should preferably be followed by one which is predominantly mental, or an operation involving standing by one involving sitting. We agree that fatigue may be one explanation, especially in those cases where long working periods are used, but it is doubtful whether the factor of fatigue operates in cases where the work is continued for only a few minutes as is the case in

most laboratory studies. At least in connection with our own experimental investigations, no work decrement occurred during a four-minute working period. This explanation is not particularly satisfactory, either, for those individuals who worked better continuously than alternately. Unfortunately, Wyatt and Fraser used comparatively few subjects for their studies, and where this is the case average results may not be very reliable.

Lossagk (13) studied continuous versus alternating activity in a work situation which he did not describe any further than to say that it consisted of loosening a screw, separating two plates, inserting a piece of cardboard and tightening it, and cutting off the edges. First the given task was performed by finishing each part, one after the other, until the whole job was completed (alternating activity); secondly, the subjects did one part at a time, let us say loosening screws, finishing this for all the pieces, then separating two plates, again finishing this task for every piece, and so on (continuous activity). Lossagk found that the average time required for doing the job ten times was 10.00 seconds for alternating and 6.16 seconds for continuous work.

In a second part of this study, the author used tasks which differ in the number of their parts, some consisting of two, others of three or four parts. In this study he found that if only a few definite movements are involved, alternating work takes less time, but if many different movements are involved continuous work is better. In other words, the greater the work unit the greater is the gain from continuous work.

The same indications appeared in Wyatt's study. There it was found that in drug packing, where comparatively greater work units existed than in some of the other work situations, such as assembling bicycle chains, for example, the results were also in favor of continuous work. This may be because the large work units or tasks are themselves varied. Another possible explanation is that the degree of integration of the activity may determine whether a given task is performed better when alternated between several parts of a given configuration or activity pattern, or when one works continuously on one part. For example, if the parts of a working situation fit together so that they form a complete unit, or a "good" configuration, and if the alternation occurs between the successive parts of this whole, then alternation should be beneficial; but if the different parts of a working situation do not fit together, or fit only very loosely, thus furnishing a bad configuration, then continuous work on one part should be more efficient than alternation between parts.

A related question is whether or not two or more kinds of material that "belong" or are related give different results under alternation than material that is unrelated. Conrad (5) presents some data on this subject, and though this is a laboratory study we shall discuss his findings here. He showed how related and unrelated material affects the ability to alternate activities. Nine subjects were required to read, for from five to ten minutes, either connected and logical passages or disconnected material which had been put together from various sources. In the middle of this period some arithmetic problems were presented to them tachistoscopically. The time needed to solve these problems under these conditions was compared with the time needed under normal conditions, when the subject was permitted to get set to add. It was found that the type of material, whether it was connected or unconnected, determined the degree of disturbance. Incoherent material that was related to arithmetic produced greater interference than did coherent but otherwise unconnected material.

Whether or not these findings apply to Lossagk's results is problematic. It is possible that the differences in the results which he obtained for activities involving various numbers of movements may be due to changing tools. Frequent changes may cause decrease in efficiency below the level of performance during continuous activity; at least Wyatt (23) found that this may be the case. In connection with train-assembling (presumably toys) he had four girls boring, screwing and adjusting wheels to the axis, either in alternation or continuously. In this activity the four girls showed a decrease of from 13 to 18 per cent in efficiency for alternation as compared with constant work, the cause for the difference in output being due to the greater number of movements and interruptions resulting from changing tools or waiting for them.

Baumberger (1) carried out a study in which the work situation in itself consisted of either continuous or alternating activity. Instead of simply comparing the average group efficiency he calculated the 100 per cent working capacity for each individual and then observed whether a given individual reached his limit more closely when working on alternating or on constant activities. This comparison was made each hour of the day with the employees working either on glass-gathering, recessing gears or operating spin tops. Glass-gathering is an operation in which the workers alternate activities, while in the other operations they perform the same task throughout the entire working shift of ten to thirteen hours. The output in the alternating task of glass-gathering shows much closer

approach to the maximum working capacity than do the operations which are continuous. In glass-gathering the average output falls only 7.7 per cent short of 100 per cent working capacity, while the other three occupations fall 31.7, 36.4, and 39.4 per cent short. The differences in these results is explained by the author as being due to differences in fatigue. Fatigue, according to Baumberger, is primarily central, the seat of fatigue being the synapse. Therefore any change in operation, although possibly involving the same muscles, would necessarily involve different synapses, and this would furnish opportunity for synapses previously used to recover from fatigue.

Since group results always cover up individual differences we cannot tell what the picture might be within a group. Again, using different subjects for the different conditions destroys any possibility of studying the consistency from situation to situation. The author's explanation in terms of fatigue is plausible, but it has some grave shortcomings, because it requires the assumption that in continuous work the same synapses are constantly employed. This is rather doubtful. Even in the most continuous activity, there is always enough variation so that the same sensory and motor cells are not always used. Considerable variability exists even in the neural structures involved in simple responses.

Vernon (20) also furnished considerable evidence for assertion that in so-called continuous work considerable change or variety of activity occurs. He reports in his investigation that the cycles of repetition do not form an unbroken series. A certain degree of variety is provided by small changes in performance, short interpolated rest pauses, and changes in posture. If we assume that varying the task so that the same synapses are not always used does reduce fatigue and therefore increases efficiency, it should be comparatively easy to test this hypothesis simply by comparing several alternating functions which involve varying degrees of difference in responses, since the wider the difference between the two alternating activities, the greater should be the increase in efficiency from alternation.

In the above study it was assumed that the differences lie in the sensory-neuromuscular system, irrespective of the external differences of the tasks. Some investigators, however, seem to believe that the differences must be differences of tasks and not of nervous function in order that increased efficiency from alternating as compared with constant work may be obtained. Thorndike (19), for example, states that the superiority observed under the varied conditions of work is most marked when the alternated tasks are widely different.

Wyatt and Fraser (23) state as one of their conclusions in connection with their study on alternating versus constant activity that high degrees of resemblance between the tasks alternated hinder one from obtaining the highest possible degree of efficiency. Alternation of dissimilar tasks, they contend, produces the greatest efficiency. However, according to Weygandt (21) dissimilarity of tasks has no consistent effect on alternation. A comparison of the results from alternation between learning syllables and learning digits, or between reading Hungarian and Latin prose, or between reading Latin and performing additions showed that dissimilar activities may produce one effect at one time and a different effect at another time.

Jersild's (8) study furnishes some data which might be regarded as meager proof that when alternation occurs between two dissimilar tasks, such as subtracting and naming antonyms, the subjects did the alternating tasks slightly better than the continuous ones. In all cases in which Jersild reports a gain from alternation, we find that it occurred when dissimilar tasks were performed. However, the difference reported between the two methods of work was only about 2 per cent, a difference too small to assure much significance. Jersild's results can merely serve as a lead for further investigations.

To summarize this section, we may say that the studies in actual work situations furnish the same evidence as the laboratory experiments, namely, that on the average, people work better when alternating between tasks than when working constantly at one task. Both the laboratory studies and these investigations show that several different factors determine the degree of efficiency in alternating tasks. Some of the factors brought out here are: (1) the type of the work; (2) the frequency of alternation; (3) the similarity or dissimilarity of the tasks to be alternated; and (4) the complexity of the tasks or the number of movements required to perform the task.

The fact that, on the average, people were more efficient when alternating than when performing constant activity was here explained largely in terms of fatigue. The fatigue is thought of by some as a central neural effect in which continuous employment of synapses during constant activity causes lowered efficiency. When alternating work is performed the synapses are used less constantly and this greater variation of the synapses increases efficiency. Along with these central fatigue effects, there is a certain amount of emphasis on local sensory and muscular fatigue and on the effects of boredom. Some investigators thought that when, in so-called constant activities, extra movements are involved, or natural variations such as variations of posture occur, greater efficiency results than

in the case of alternation between two tasks. Again, it was suggested that the degree of integration, or whether the task formed a "good" or a "bad" configuration, determined whether people work better when alternating than when performing a continuous task.

Our next steps are first, to discuss the findings of the investigations on individual differences in ability to alternate, and next, to show that the findings and explanations from the group studies complement those from the studies on individual differences.

Individual Differences in Alternating

The problem of individual differences in alternating between tasks has been approached entirely apart from the general problem of ability to alternate. All the investigations in this field are based on some notion of "perseveration". The degree to which perseveration is present or absent supposedly determines an individual's ability to alternate. Now, if perseveration is the thing which causes a person either to be a good or a poor alternator, we may well ask ourselves, "What is this thing 'perseveration'?" If one analyzes the works on perseveration it becomes evident that considerable confusion exists as to the meaning of the term. It has been used to describe several psychological phenomena. Sometimes it is employed to denote the occurrence of spontaneous "ideas"; at other times it is used to explain the persistence of after-effects or any kind of neuromuscular activity beyond the duration of the appropriate stimuli. At still other times the term perseveration is used to explain the inhibitive effects of an activity upon a subsequent activity. It must also be pointed out that the concept has been made an entity, and used as if it were a name of a something which makes people work better alternately or continuously, depending upon the amount of that something present. Different writers have used the concept in so many different ways that it is frequently very difficult to decide what a given author means when he talks about perseveration.

If, however, we attempt to sum up, as far as possible, what is meant by perseveration in connection with studies of alternation, the idea is something like this: Every organism manifests a tendency to have its activities persist, and the degree of this persistence or perseverative tendency varies from person to person, thus producing differences in the ease with which these persons can break up a given "set" and change to new activities. Whatever activity a person engages in will hinder subsequent activities to greater or less degree, and the degree of this hindrance will depend upon the degree of perseverative activity. According to this point of view people should

work better at continuous tasks, although the investigations on individual differences in ability to alternate indicate that there are people who work better when alternating.

In order to explain this latter fact according to the perseveration hypothesis, it was maintained that these people possess only minimal perseveration or none at all. It is at this point that the apparent contradictions between the findings from the group studies on alternating versus constant activity and those on individual differences in alternating ability appear. As we have seen in the first part of our study, most of the people, on the average, work better alternately than continuously. This phenomenon was explained variously as being due to the fact that in alternation either less local fatigue, or less refractoriness, or less monotony, or less persistence of specific responses seemed to interfere with the performance than in continuous work.

It seems probable that individual differences in ability to alternate should not be explained in terms of perseveration alone, but in terms of some or all of these factors brought out in connection with the group studies, namely, the kind of task to be alternated, the frequency of alternation, the similarity or dissimilarity of the tasks, and their integration or quality of configuration. All these factors are seemingly important in determining individual differences in ability to alternate.

But instead of entering any further into theoretical discussions, let us now analyze some of the studies which bear upon the problem of individual differences in alternating activities. These studies can be separated into three groups. Those in the first group bear upon the problem only indirectly. In them the investigators tried to establish the fact that perseveration is a general "trait" in the same sense as general intelligence. Because in these investigations perseveration was measured in terms of effect of one activity upon another subsequent activity, and because it was considered a general "trait", differing in degree in different individuals, the investigators who studied the problem of individual differences in alternating ability considered perseveration to be the determining factor. We shall analyze some studies from this group in order to show that perseveration, as apparent in effects from alternating tasks, is not as general as it might seem at first sight, and that merely proving the existence of a general factor of perseveration does not prove that this is the sole factor in determining ability to alternate. The studies of the second group test the relation between intelligence and ability to alternate. Into the third group fall those studies which are based

on the assumption that perseveration is a general factor which controls ability to alternate tasks and which, accordingly, separates people into "types" of good versus poor alternators.

Perseveration as a General Factor and as Cause of Individual Differences in Ability to Alternate. Jones (9) was one of the first who tried to show that "perseveration" has a functional unity and that it is a general factor just like intelligence. Seventy-seven children, averaging about twelve years of age, were given four perseveration tests: (1) writing S's, first in the usual way, then as they appeared in a mirror; (2) writing digits in the usual way and then with the stroke backwards; (3) mirror drawing; (4) copying prose in the usual way, then without dotting the i's or crossing the t's. The results from these perseveration tests gave an average intercorrelation of .49.

In order to see whether this group of tests could be accounted for by "G" and specifics, or whether these specific factors form a group factor or factors, Spearman (18) applied his "reference test" method to part of Jones' data. According to Spearman, group factors over and above "G" may be detected by this method. The mathematical procedure is described in the appendix of *The Abilities of Man* (p. xxiii). Practically speaking, the author first measures the amount of "G" in a given test or in each of a series of tests, and then by means of partial correlation determines the amount of relation that is left over in any two given tests. If we assume that the specifics of any two tests are uncorrelated, and if "G" has been eliminated, then, if any further correlations are found they may be attributed to the overlapping of group factors. Spearman reports one tetrad difference from two "reference abilities" and two perseveration tests. This tetrad difference equals .399 with a probable error of $\pm$.048. The difference is more than seven times its probable error and thus, according to the author, is seen to be far beyond what can occur by mere sampling. Hence, he concludes, "The evidence for some group factor or factors pervading these tests of perseveration leaves nothing to be desired" (18, p. 297). The specific correlation between the two perseveration values is .600.

Granted that according to this analysis, a group factor between the two perseveration values is present, these findings from one tetrad difference do not permit one to draw any generalizations. Why does the report give only one tetrad difference when three were possible, and specific correlations for only two tests when four tests were used and thus six combinations could have been made? All the tetrad differences from these tests should be considered in reaching any

conclusions even for these two tests. Furthermore, Spearman discusses Jones' results in such a way as to convey the impression that the finding of this group factor proves that it is perseveration. If we study the logic of the tetrad theory, however, all that may be concluded from the analysis is that the tetrad equation is not satisfied, and that the tetrad difference is more than five times its probable error, which would indicate that the two-factor hypothesis does not hold, and that there is not only one common factor to account for the intercorrelations, but that one or more group factors are present. These results do not show that there is only one additional factor and that this factor is perseveration. The author admits himself that Jones' results do not show that the additional factor "essentially consists in perseveration"; yet he decidedly treats the results in that way.

Pinard (16) applied four tests of perseveration to 194 "difficult" children and 116 adults. These tests were: (1) an inverted S test, in which the subjects wrote S's in the usual way for thirty seconds, and then inverted them for the same length of time, and finally wrote the two kinds of S's in alternation. (2) The second test consisted of triangle drawing, ($\triangle \triangledown$). (3) In a third test the subject alternated between writing *abcdefg* and 234567. (4) As a final test the subject had to write B,C,D,R,F, in normal and mirror script. With the exception of test two, these tests were the same as those used by Jones. The tests were applied five times after five minutes of practice and perseveration was measured by subtracting the total number of correct items written in the second 2 min. from the total in the first 2 min. The results from the 194 children showed an average intercorrelation of .34. The tetrad differences for the above correlations were worked out. The mean deviation of tetrad from zero was found to be .0263, and the probable error .02104. The theoretical probable error which one should expect to get according to Spearman's formula was .0217. The probable errors are thus almost the same, and approach the mean tetrad, which seems to indicate that one general factor and some specific ones are sufficient to explain the intercorrelations of the four tests. This general factor the author calls perseveration.

But why call it so? Since Pinard's results satisfy the conditions of the two-factor theorem, according to Spearman's interpretation of Jones' results, it should really be called "G". But if Pinard considers the two-factor theorem a purely mathematical one, as Holzinger and Spearman suggest, though the latter does not always give that impression in his book, then he may call his general factor

perseveration or anything else he desires. Whether other psychologists will agree with this name, however, is another problem, and one which is much harder to solve than the first. This approach raises an additional difficulty for Spearman's treatment of Jones' data. Since both Jones and Pinard used practically the same tests and the same procedure, why did Pinard find that the two-factor theory is satisfied, while Spearman shows that it does not hold, but that there are some additional factors? Will these opposed results both make a case for perseveration?

Lankes (12) performed the most extensive study on perseveration. He was one of the first to intercorrelate the scores from many varied tests. He was particularly interested in discovering whether different forms of perseveration, such as persistence of after-sensations, alternating effects, and perseverations of "ideas" or experiences in everyday life intercorrelate. His tests were so constructed as to bring out the effects from these forms of perseveration. One test consisted of the "natural rate of tapping". The subject was told "simply to move the finger at its own rate, just as he feels it natural to himself at the time". Another test consisted of letter-writing. Here, the subjects wrote six letters, let us say *lmnopq*. First they had to write these letters many times in alphabetical order, and then a few times in the reversed order. A cancellation test was used as a third test. In that, first one letter then another had to be cancelled. A fourth test was a color-fusion test. In a fifth test the subject saw drawings which were exposed for six seconds either one after the other or in pairs, in which each drawing was exposed alternately for two seconds, first the left, then the right, etc. The sixth test consisted of narratives which were read to the subject. First one story was read whereupon the subject had to answer questions; then immediately following, another story was read and questions given. For a seventh test Lankes required his subject to free-associate to a series of stimulus words. Finally, as an eighth test the subject had to write brief essays on different subjects. About four minutes of writing was required for one essay, of which there were five at each sitting.

Perseveration was measured in two ways: (1) directly, by noting the slowness of the rate of tapping and the rate at which color-fusion occurred; (2) by finding the degree of hindrance which the perseveration effect of a past activity caused on a new one. This last was expressed by a quotient between the average of several days of work without alternation and the average when alternating was done. Each test was performed for thirty seconds at a time, and four such

trials separated by one minute were made per day. In addition to all these tests, the eighty subjects had to answer a lengthy questionnaire constructed for the purpose of bringing out perseveration tendencies in everyday life, or the spontaneous occurrence of past "ideas".

The intercorrelations of the results of the perseveration tests and the questionnaire range from —.05 to .50, with an average r of .41. Although no tetrad analysis had been made or at least is not mentioned in the report, the author believes that these correlations should be attributed to a perseveration factor which is general in all the tests. He believes that this factor is "P" rather than "G" because, as he says, the tests had been so constructed as to eliminate "G".

Even these few studies are sufficient to show that the results are very equivocal, and that they furnish no conclusive evidence that perseveration is a general factor, comparable to general intelligence. In these studies perseveration may only be considered general in the sense of being a common factor to a series of tests, and not general in the sense of operating in all activities.

Considerable confusion exists between Spearman's method of analyzing the intercorrelations of a group of tests for a common factor and his theory that perseveration is a general factor. Even though it should be possible to demonstrate that the intercorrelations from a series of perseveration tests can have a common factor, this does not prove that these correlations can be accounted for by only one factor, for there might be still other factors present. Again, the demonstration of a general factor "P" does not prove that this mysterious something is a native quality of the nervous system and that it determines an individual's ability to alternate. If we keep these distinctions in mind we will be better able to realize that the previous studies provide no evidence that the degree of perseveration determines an individual's ability to alternate. That other factors may be equally important is shown in an experimental investigation of our own.

Relation of Ability to Alternate and Intelligence. There is a difference of opinion among investigators as to whether ability to alternate is due to perseveration or merely to general intelligence. Spearman (**18**, p. 306) thinks that by means of his two-factor analysis he is able to show that perseveration, or what he calls inertia, or "P", is a general factor independent of intelligence. He believes that "G" and "P" vary independently of each other but that perhaps they are different aspects of the same thing, the mental energy; as "G" measures its quantity, so "P" may express its degree of inertia. Jersild (**8**) argues very fervently that the ability to shift is due to general ability. His results from the Otis Self-Administer-

ing test, which he gave to one group of his subjects and from the Army Alpha test which he gave to another group, show that on the average the correlation between intelligence scores and scores in alternating work is higher than the correlation between intelligence scores and scores in continuous work. The Otis tests gave lower correlations (from .25 to .52) than did the Army Alpha test (from .39 to .67). The tests showing highest shift loss show the highest correlation between shift and intelligence. The author explains the correlation of the results from the perseveration tests and those from the intelligence tests by a third common cause, general ability. Thus, according to this author, it is general intelligence which enables a person to alternate between tasks.

Zillig (25) made a comparison of her results of "Umstellbarkeit" on arithmetic with the class marks which the pupils received in that subject, and found results opposite to those obtained by Jersild, namely, that the poor students are better in alternating, while the good students are the reverse. The significance of these results may be questioned, since class marks are not always a very reliable measure of intelligence. Our present evidence regarding the relationship between ability to alternate and intelligence is not very convincing in either direction.

Types of Alternators. Mary Zillig (25) attempted to establish perseveration as the general cause of a person's ability to alternate from task to task, and also to develop a system of types of alternators. Since people show varied degrees of perseveration, Zillig believes that definite alternating types can be found and that there are good and poor "Umstellers". There should be people, according to Zillig, who are much more efficient when they can work continuously on any given task while there are others who work better when they can change from task to task. Those who work better continuously she calls poor "Umstellers" on the assumption that they cannot change set quickly. The others, who supposedly can change set easily, are the good "Umstellers". Whether this "Umstellerbarkeit" is a distinct ability to change set or whether it is merely the result of the particular degree of a person's perseverative activity, Zillig does not state. But one may infer from her way of treating her material and from her references to the work on perseveration that she probably holds the latter point of view.

Whatever Zillig's point of view may be, she thinks that she has established the existence of these two types of "Umstellers" by a series of experimental tests. She had a total of 46 subjects of which 24 served as subjects in all the experiments. These subjects were

third and fourth grade pupils. They had to perform five different tasks, either working continuously or alternating between two tasks. First the pupil's attention was tested by having him cancel the digit one in a cancellation test five times, and then cancel another digit twice, etc., alternating between the two for ten minutes. The results from this procedure were then compared with the results of ten minutes of cancelling i's continuously. Next, Zillig tested a pupil's ability to change between oral addition and subtraction. In this situation the child either added continuously or changed between eight additions and three subtractions. In still another test, ability to alternate between motor tasks was studied by having the subject write ◡ ten times either continuously or alternating with writing ◠ five times. "Umstellbarkeit" in manual work was tested with ball-tossing. Here the subject had to catch a ball alternately overhead and below the level of the head. Finally, there was a test in which cards had to be sorted according to different signs and combinations. In each test a subject was given five seconds for one response. Then the subjects were ranked for every continuous and every alternating task. The quality of the work was measured by counting the number of errors produced either by omission or by wrong work; when the subjects did not change at all two errors were counted.

The report of the experiment in which the author attempted to classify her subjects gives data from 23 subjects, 11 of whom, on the average, worked better alternating, and 7 of whom did better at continuous work. With five subjects it made no difference either way. However, there is not 100 per cent consistency, for a given subject frequently does not run true to type. The kind of task also determined different amounts of "Umstellbarkeit". Further, for some individuals the fact that they knew or did not know that they had to change also helped to determine whether they were good or poor "Umstellers".

In spite of the fact that Zillig's data showed many exceptions to the rule, and that she admits these exceptions, she concludes her study as if she had proved the existence of two general types of "Umstellers". She puts her subjects into fast categories just as if her classifications would hold for all possible work situations, regardless of the nature and the kind of material, the period of work, or the kind of tasks involved.

This problem of "type" has important practical bearings. If it should be possible to prove that some people are consistently more efficient when performing continuous as compared with alternating

activities, while others show just the reverse tendency, then the problem of efficiency in repetitive work in factories and offices might be solved by choosing the right man for the right place. Furthermore, if two distinct types of people should be found, those with poor and those with good alternating ability, the schools ought to adopt teaching methods which are adjusted to the needs of both "types" of students.

In Germany, some scholars in the field of psycho-technology have applied this typology of "Umstellbarkeit" to aptitude analysis and employment selection. For example, von Kuegelgen (10) in connection with the construction of a battery of tests for the selection of office employees, used an "Umstellbarkeit" test which he had built upon the principle of those used by Zillig.

Kuegelgen believes that if a subject obtains a good "Umstellbarkeit" score on his test it may be relied upon that this particular employee will be adept in alternating from one type of office work to another. However, in most cases the actual work of an office is much more complex than the tests and not at all similar to the task by which the subject's "Umstellbarkeit" has been determined. Therefore, such a test procedure is valid only if we can assume that "Umstellbarkeit" is a general factor in a person's make-up. If a person's ability to alternate between tasks shows little consistency from situation to situation or from task to task, then, such a generalization on the basis of a single test may lead to grave mistakes. If "Umstellbarkeit" cannot be explained by a general factor, different tests specific to particular situations must be employed. An individual is then a poor or good "Umsteller" for such and such a test, but not in a general way.

In summarizing the results of the studies on individual differences in ability to alternate, we may say that most of these studies are based on the assumption that an individual's ability is determined by his degree of "perseveration", which is considered to be a general factor similar to general intelligence. It was shown that "perseveration" is not proved to be a common factor in such a general way, although some series of tests have been found to possess a common factor. The presence of a common factor does not prove that no other common or group factors may operate; nor does its statistical demonstration prove that it is perseveration and that this is an innate quality of the organism. The results of the "perseveration" studies in no way prove that "perseveration" is the only factor in determining an individual's ability to alternate. Perseveration may contribute to it but other factors may be just as important.

The few attempts to compare alternating ability and intelligence give little conclusive evidence for or against the idea that this ability is independent of intelligence.

Again, no crucial test has been made of whether or not some people are poor and others are good alternators. There is very little justification for the assumption that this ability is very constant from situation to situation regardless of the task.

Résumé of the Present Status of the Problem of Alternating Activities

The findings from our analysis of the present status of alternating ability in general and individual differences in particular may be summarized as follows:

(1) The problem of individual differences in ability to alternate activities is closely linked up with the problem of alternation in general. Therefore the group studies on alternation versus constant activity and those on individual differences *per se* must be viewed together. Both groups of studies furnish valuable data for the understanding of individual differences.

(2) The findings from the group studies, whether performed in the laboratory or in actual work situations, give results, on the average, in favor of alternation. Hence we may conclude that most people usually work better when alternating between tasks than when working constantly at one. But, the effects from alternation seem to vary according to the working conditions.

(3) Some of the conditions or variables which aid in determining whether constant or alternating activity is more effective are: the kinds of tasks alternated, whether arithmetic or soap-wrapping, for example; the similarity or dissimilarity of the tasks alternated; the frequency of alternation; the condition of the working situations, that is, whether its parts are all integrated and form, so to speak, a good configuration; and the difficulty of the tasks alternated.

(4) The superior efficiency from alternation was explained in terms of absence of fatigue, especially local fatigue, according to which continuous employment of the same neuromuscular parts during constant activity causes lowered efficiency. When alternating work is performed the synapses are used less constantly, and thus by greater variation of synapses efficiency is increased. In other situations, superior efficiency during alternating work has been explained as being due to less refractory phase effect during alternation than during constant activity; or again it was thought that alternation brought about superior efficiency because of absence of

interference of specific responses which lower a subject's efficiency during constant work. Explanations have also been suggested in terms of boredom and of energizing work-sets.

(5) Inferior average performances for alternating as against continuous work were sometimes explained in terms of "perseveration", according to which hypothesis persistence of a general set interferes with the process of changing from task to task. Some investigators thought that when, in so-called constant activities, extra movements, or natural variations, such as variations of posture and short rest pauses are involved, greater efficiency results than when there is alternation between two tasks.

(6) The analysis of the studies dealing more or less directly with individual differences brings out the fact that these differences in alternating are not necessarily due to individual differences in "perseveration". Such differences may contribute to ability to alternate, but no convincing evidence was found that they are the only or even principal causal factors.

(7) Comparisons of alternating ability and general intelligence, too, gave comparatively few clear-cut results in favor of the idea that ability to alternate is due to intelligence; nor has it been proved that ability to alternate is a special ability.

(8) The problem which has most practical bearing, namely, that of whether some people consistently work better when alternating, while others consistently do better at constant activities, has been least investigated, and the few studies which have been made furnish little convincing evidence for the assumption of an alternating "trait". No crucial test has been performed which proves that some people are good and others are poor alternators. An individual might have the tendency to be either a good or a poor alternator, but there is little justification for any assumption that this ability is constant from situation to situation, regardless of the task.

Before any general conclusions can be drawn as to whether or not there are alternating "types", one should compare the results from alternating versus continuous work in many varied situations, in order to see whether people differ consistently in doing alternating versus constant tasks. Their ability to alternate between predominantly manual or mental tasks should be compared, or different types of mental or motor tasks with efficiency when any of these tasks are performed constantly. Again their efficiency should be compared for alternating and constant activity under various frequencies of alternation, or for alternation between several tasks, or for various alternating conditions. One might go on naming possible

variations, but the above-mentioned ones suffice to indicate that, ideally, before an individual can be called either a good or a poor alternator he should be consistent in type in every possible situation or at least in a great variety of situations and tasks. It is admitted that it is practically impossible to test a person's ability to alternate in every possible situation, or in every possible combination of tasks. However, it is possible to choose some variables which, from the analyses of the previous studies on alternation versus constant activity, appear important in determining whether on the average people perform better alternately or constantly, and in testing a group of subjects to see whether they show a great consistency from task to task or from situation to situation. Then it may be hoped that other situations will give this same consistency.

It was with this point of view that the following investigation (3) was planned: Fifty-one college students worked continuously on addition, subtraction, supplying verbs or adjectives to nouns, drawing triangles point up or point down, and forming two kinds of figures from groups of six dots. They then alternated between two of these activities. The tests were paired to form either similar or dissimilar activities, such as addition and subtraction, or addition and drawing triangles. There were four pairs of similar and four pairs of dissimilar activities. Two frequencies of change were tested by having the subjects alternate every minute or every ten seconds. It was found that people differ greatly in their ability to alternate activities, and they show little consistency from task to task. The subjects could not be considered good or poor alternators in a general sense. They showed little consistency in their ability to do varied as against constant tasks. A given individual may be a good alternator in one situation and poor in another one. Ability to alternate was found to depend upon the type of the tasks alternated, and upon whether the shift occurred between similar or dissimilar tasks. Even by adopting a very loose criterion of type, very little tendency for a subject to be either a good or bad alternator was observed. When using similar tasks 38 out of 51 subjects worked better continuously; while, when they alternated dissimilar tasks, 47 worked better at alternating, and only 4 at constant activity. Varying the frequency of alternation from one minute to ten seconds produced little difference in results for alternating versus constant activity.

BIBLIOGRAPHY

1. BAUMBERGER, J., The Effect of Change of Work. *J. Indus. Hyg.*, 1923, **5**, 245–252.

2. BILLS, A. G., and MCTEER, W., Transfer of Fatigue and Identical Elements. *J. Exper. Psychol.*, 1932, **15**, 23–26.
3. BURRI, C., Individual Differences in Ability to Alternate Activities. Doctor's dissertation, June, 1933. To be published in *J. Exper. Psychol.*
4. CHAPMAN, S. C., The Effect of Rapid Change of Work on the Rate of Performance. *J. Exper. Psychol.*, 1917, **2**, 165–170.
5. CONRAD, W., Einstellung und Arbeits Wechsel als Pädagog. u. Allgem. Psycholog. Probleme. *Arch. f. d. ges. Psychol.*, 1915, **34**, 317–514.
6. DASHIELL, J. T., *Fundamentals of Psychology*. New York: The Houghton Mifflin Co., 1928, Chap. 10, 281–282.
7. HOLZINGER, K. J., *Statistical Résumé of the Spearman Two-Factor Theory*. Chicago: The University of Chicago Press, 1930. Pp. 43.
8. JERSILD, A. T., Mental Set and Shift. *Arch. of Psychol.*, 1927, **14**, No. 89. Pp. 82.
9. JONES, W., *Perseveration*. Report, Brit. Assoc. for Adv. of Science, 1915. London: John Murray, 1916, 698–699.
10. VON KUEGELGEN, G. W., Eignungsprüfung für den Kaufmann Büreau-angestellten. *Psychotechn. Zsch.*, 1932, **8**, 70–84.
11. VON KRIES, J., Über die Natur gewisser Gehirnzustände. *Zschr. f. Psychol.*, 1875, **8**, 1–32.
12. LANKES, W., Perseveration. *Brit. J. Psychol.*, 1915, **7**, 347–419.
13. LOSSAGK, H., Werkstück und Werkzeuggriffe bei verschiedenen Arbeitsweisen. *Indus. Psychotechn.*, 1931, **8**, 289–299.
14. MILES, G. H., and SKILBECK, O., An Experiment of Change of Work. *J. Nat'l Inst. Indus. Psychol.*, 1923, **1**, 236–240.
15. MÜLLER, G. E., und PILZECKER, A., Exper. Beiträge zur Lehre vom Gedächtnis. *Zsch. f. Psychol.*, Erg. Bd., 1900, **1**, 58–60.
16. PINARD, J. W., Tests of Perseveration: (Their Relation to Character). *Brit. J. Psychol.*, 1932, **23**, 5–19.
17. ROBINSON, E. S., and BILLS, A. G., Two Factors in the Work Decrement. *J. Exper. Psychol.*, 1926, **9**, 415–443.
18. SPEARMAN, C., *The Abilities of Man*. New York: The Macmillan Co., 1927, Chap. 17, 291–307.
19. THORNDIKE, E. L., *Educational Psychology*. New York: Columbia University, 1914, **3**, 135–136.
20. VERNON, H. M., *On the Extent and Effect of Variety in Repetitive Work*. Indus. Fatigue Research Board Report 26, 1924, 1–12.
21. WEYGANDT, W., Ueber den Einfluss des Arbeitswechsels auf fortlaufende Arbeit. *Psychol. Arbeit.*, 1897, **2**(1), 118–202.
22. WYATT, S., *The Effect of Change in Activity*. Indus. Fatigue Research Board Report 26, 1924, 13–37.
23. WYATT, S., and FRASER, J. A., *The Comparative Effects of Variety and Uniformity in Work*. Indus. Fatigue Research Board Report 52, 1928, p. 36.
24. ZIEHEN, T., Die Ideenassoziation des Kindes. *Sammlung v. Abhandlungen a. d. Gebiete d. Pädagog. Psychol.*, 1898, **1**, No. 6, 66.
25. ZILLIG, M., Experimentelle Untersuchung über Umstellbarkeit. *Zsch. f. Psychol.*, 1925, **95**, 274–315.

PSYCHOPHYSICAL MEASUREMENT METHODS

BY FRANCIS W. IRWIN
University of Pennsylvania

During the nine years[1] since the latest review of psychophysical measurement methods by Rich (120), the growth of interest in this field has suffered no relapse. New methods, refinements and generalizations of old ones, and applications of psychophysical principles to an ever widening range of problems, both theoretical and practical, have marked the period. We shall take up in order the following topics: The Comparison of Psychophysical Methods, The Method of Single Stimuli, Categories of Judgment, Special Influences upon the Judgment, Statistical Considerations, General and Theoretical Considerations, and New Applications.

Comparison of Psychophysical Methods. Among the major papers on the comparison of psychophysical methods was that of Kellogg (83) in which an effort was made to determine the relationship between the methods of average error and constant stimuli. Kellogg wished to set up equations whereby the values obtained by one method could be related to those of the other, and to determine which method was the 'better' by the criteria of reliability, accuracy, magnitude of sensitivity values yielded, and applicability to the study of Weber's law. Assuming that the method of average error gave no limen, Kellogg treated his results in terms of the variable error (VE) and constant error (CE) provided by each method, the intermediate category of judgment having been excluded from the method of constant stimuli. In general the VE's and CE's of constant stimuli were larger than those of average error, the differences

[1] The present review is intended to cover the period from 1925 to 1933, inclusive. A few incursions into 1934 were made when it seemed possible to make more final the discussion of a particular subject. It is becoming increasingly difficult to survey within one paper of reasonable length the whole territory once included by reviews under this title; hence the writer has omitted from consideration such work as that on the range of apprehension (for which see Dallenbach, 29, 30, 31) and the conditioned reflex technique as used in the measurement of sensitivity (see, *e.g.*, Razran, 118), and has not attempted to include in either text or bibliography the enormous literature on the study of the psychophysical constants and the validity of Weber's law for the various sense modalities (most of which is reviewed under various titles in this journal) except where such material seems relevant to methodology.

in VE usually being statistically significant. On the other hand, the rank order of observers was found to be more consistent and the variability of the VE less, in the case of the method of constant stimuli. Only approximate predictions could be made from equations relating the VE's of the two methods.

In an evaluation of Kellogg's work, Urban (162) stressed the necessity of distinguishing between the 'material' and the 'formal' conditions of a method, the former being all factors influencing the probabilities of judgment, and the latter consisting of the conditions affecting the calculation and evaluation of the numbers of relative frequency. A valid comparison demanded the presence of experimental results which could be calculated by both methods, since only under these conditions would the material conditions be the same. The unsatisfactory result of Kellogg's attempt to predict from one method to the other Urban ascribed to the failure to keep the material conditions identical and to the exclusion of equal judgments from the method of constant stimuli, the latter being important because, as Urban believed, it eliminated the only psychologically common factor in the two methods. He agreed with Kellogg that the determination of the 'better' of two methods should involve the setting up of equations relating them, but considered irrelevant the magnitude of the index of sensitivity and the applicability of a method to the study of Weber's law. Here it seems to the reviewer that the term 'better' necessarily refers to a particular purpose and that Urban is correct to the extent that the purpose of a psychophysical method is solely that of determining an index of sensitivity. But can it not be argued that the finer the method of sensitivity the closer it is to the 'true' measure, *i.e.*, the measure as properly defined?

Kellogg's assertion that no threshold value was yielded by the method of average error led Urban to analyze in another paper (161) the relation between this method and that of just noticeable differences. Assuming that in the method of average error there was a finite interval in which the psychometric function of equal judgments had the value of unity or closely approached this value, he showed that the probabilities of individual judgments in the method of average error were the same as those which gave the upper and lower limits of the I of U in the method of just noticeable differences, and that the average error corresponded to the midpoint of the I of U, or the PSE of the method of just noticeable differences. The Gaussian coefficient of precision, h, was inversely proportional to the I of U, the relation being $h=\frac{2}{\sqrt{\pi}\,(S_1-S_2)}$ where S_1 and S_2 were the upper

and lower limens. These results should be of the utmost importance for the understanding of the method of average error, with its peculiar advantages in obtaining clinical and anthropometric measures where unpracticed and, often, uninterested Ss must be employed.

Thurstone (**146**) studied the method of equal-appearing intervals by means of a variation of Sanford's experiment, using instead of weighted envelopes a series of cards more or less densely covered with irregularly arranged dots. One hundred and one Ss sorted the cards into 10 piles evenly graduated as to the apparent density of the dots. The median interval into which each card was sorted was plotted against the logarithm of the number of dots on the card; a straight line resulted. Thurstone concluded that this constituted "final proof that Fechner's law holds for the stimuli of the present experiment" (p. 220). He criticized the method, however, on the grounds that end-effects were shown in the distributions at some 'sensation steps', the distributions being cut off and non-normal, and that the determinations were vitally affected by the distribution of stimulus values used in the experiment. These data were considered critically by Urban (**164, 165, 166, 167**). He disagreed with Thurstone's assertion that the distributions were not normal, saying that they appeared to be normal though it was true that many were cut off at the ends due to the procedure of the experiment. When the constants of the psychometric functions for each 'pile' or column of Thurstone's data were calculated according to the $\Phi(\gamma)$ hypothesis, the I of U's were found to increase irregularly from one end to the other of the intervals 2–9. The PSE of each column was held by Urban to be the correct measure of the stimulus value appearing equal to the standard stimulus (in this case, the 'subjective standard'). These PSE values differed considerably from the *averages* calculated according to the usual procedure in this method. The values critical for the study of Weber's law were, according to Urban, the ratios of the I of U to the PSE, which showed considerable variation. Those values critical for Fechner's law were the ratios of each PSE to that of the succeeding interval, and in Thurstone's results they were very nearly constant. On the basis of this experiment Urban asserted the mutual independence of Weber's and Fechner's laws.

Comparing the method of paired comparison, the graphic rating scale, and two variations of the order of merit method, Tinker (**156**) found high intercorrelations. The reliability of the methods varied with the kind of material employed.

Culler (**25**) advocated the use of a variable standard stimulus in the method of constant stimuli. As advantages of this procedure he cited the wider range of stimulus differences made available, the relatively larger number of small differences in the series, and the elimination of 'absolute impression'. He believed the practical value of this procedure demonstrated by his work on temperature sensation (**22**).

Culler (**25**) also defended the use of the 'stimulus attitude' in psychometric work. Analysis of the meaning of the term 'stimulus error' showed, he thought, that the term itself was not apt. Use of the stimulus attitude is valuable because it comes naturally to the S and because it leads to greater precision as shown by several investigators. It is based, according to Culler, upon a complex criterion with more cues at hand for judging; an error is introduced only when the criterion contains cues irrelevant to the thing measured.

Shaad and Helson (**129**) studied the possibility of the repetition of stimuli in the method of constant stimuli as a time-saving device. Although the precision of this method seemed slightly lower than that of the usual chance order of stimuli, the authors concluded that the method could be used with profit. Further work appears advisable.

A new and interesting approach to the comparison of methods was opened by Linder (**97**); since it is purely statistical it will be discussed below in the section on statistical considerations.

As one familiarizes himself with the work done on the psychophysical methods in this and the preceding periods he is likely to sympathize with the point of view expressed by Bressler (**14**): "It may safely be said that more energy has been expended in standardizing the different methods and in discovering their comparative utility than in utilizing them. And still none of them has been completely standardized" (p. 7). We might add that too often the intrinsic interest of a psychophysical study is in inverse proportion to the thoroughness of the methodological precautions observed, so that often one is immediately interested by the statement of a problem only to find little evidence of knowledge of the refinements of method made available by years of development in the field. Unfamiliarity with work in foreign languages is often apparent on both sides of the Atlantic. In general, the choice of a method still seems to be very much a matter of the experimenter's convenience and temperament.

The Method of Single Stimuli. The phenomenon of 'absolute impression' has been recognized since the early work of Martin and

Müller[2]; recently it has led to the development of a new method, that of 'absolute judgment' or 'single stimuli', at the hands of Truman and Wever (**158**) and Wever and Zener (**176**).[3] The latter authors compared a 'normal' series of lifted weight judgments with series in which no standard weight was used, judgments in the latter series being given in the categories '1', '2', and '3' for the light to heavy impressions in one case, and 'light', 'medium', and 'heavy' in another. They found it well to require their Ss to begin giving judgments immediately. Psychometric functions of the usual form were obtained in the 'absolute' series; using the criterion $\Sigma(p'-p)^2$ the fits were at least as good as those of the relative series. The precision appeared not to suffer from the omission of the standard stimulus. An analogy of the usual negative time error appeared, in that the 'PSE' was lower than the middle stimulus value. A larger number of equal judgments were given in the absolute method with correspondingly greater intervals of uncertainty, but the authors believed this was related to the Ss' attitude and was not an essential characteristic of the method. They drew the general conclusion that the new method was practicable as a true psychophysical method. Since the peculiarities of the method lead to certain terminological difficulties Volkmann (**168**) published a note in which terms analogous to those of constant stimuli were agreed upon by a number of interested experimenters. It is suggested that they be accepted by all those who discuss the new method.

Believing that Gahagan (**53**), who studied the possibility of using several comparison stimuli with one standard, and Wever and Zener had not sufficiently controlled the practice effect, Fernberger (**44**) obtained 5 series of 50 judgments on each comparison stimulus by the usual method and followed this by an equal number of judgments by the method of single stimuli. The precision was about the same in the two series. The 'central interval' (CI) of single stimuli was considerably greater than the I of U of constant stimuli, due to the larger number of 'intermediate' judgments in the single stimuli series. Six of the eight Ss had CI's above the 70 per cent point in the distribution of 58 previous Ss whose I of U's had been obtained

[2] Martin, L. J., and Müller, G. E., *Zur Analyse der Unterschiedsempfindlichkeit.* Leipzig, 1899.

[3] To Truman and Wever goes the credit for establishing experimentally the availability of the 'absolute' factor for use in a psychophysical method; Wever and Zener have priority in the complete elimination of the standard stimulus and the thorough development of the method. Various procedures in which a standard stimulus was not used, but dealing chiefly with qualitative material, have long been familiar; see Wever and Zener (**176**) and Beebe-Center (**6**).

by the customary method. The PS ('point of symmetry') was invariably lower than the PSE, confirming Wever and Zener's findings. (On this point see also Pratt, 117.) Practice appeared to have similar effects in the two methods; in the method of single stimuli practice was concluded to consist largely of the forming of definitions of the judgment categories. Some indication was found that relative series following single stimuli series gave I of U's larger than usual. This was investigated by Davies and Babb (33) whose results were not clear-cut, but in general favored Fernberger's finding. The tendency seemed not to be due to the carrying over, from single stimuli to the relative series, of an attitude toward giving a larger number of intermediate judgments.

Progress with the method of single stimuli has warranted the admission of the method to the consideration of anyone intending to make psychophysical measurements. Very striking is the repeated finding of comparable precision with the older method. It must be remembered, however, that as shown by Culler (25) and repeated by Kellogg (84), the presence of a large number of equal judgments tends to heighten the precision; and the method of single stimuli is characterized by its tendency to lead to the giving of such judgments. It would seem possible, therefore, that the method of single stimuli ought to give consistently *greater* precision than the usual method unless there is some factor tending to counteract the influence of the intermediate judgments. The reviewer ventures the consideration that such a factor might be found in the difficulty of judging in the absence of the standard stimulus. If the CI is to be used as the measure of sensitivity the method of single stimuli needs further study, the CI having been found to be relatively large whereas the smallest possible value has some claim to consideration as the best measure of sensitivity (*cf.* Kellogg, 83, and our discussion above). If, as Wever and Zener suggest, the difference is merely one of attitude, then an experiment in which attitude was carefully controlled so as to give relatively few intermediate judgments ought to give both precision and CI comparable to the corresponding measures of constant stimuli. Until this has been shown to be so, and until experiments have been performed in other fields than lifted weights (particularly vision)—see, *e.g.*, Fry, Haupt, and Wartena (52) on the two-point limen—a great deal of caution will be necessary in interpreting the single stimuli results as entirely comparable to those of constant stimuli.

Categories of Judgment. The problem, especially annoying in the method of constant stimuli, as to what and how many categories

of judgment should be allowed, has been the subject of frequent discussion. Much distrust has surrounded the category 'equal' because it seems particularly susceptible to influences irrelevant to the measurement of sensitivity, such as practice, attitude, and temperament. Fernberger (41) contended that an intermediate category could and ought to be used, with adequate control of subjective factors. Culler (25) also defended it, on the grounds that it made judgments more precise, that it avoided compromises on the part of the subject, and that it helped the subject keep a check on his judging. Culler held, however, that the category 'doubtful' was preferable to 'equal'.

Believing that the variability of equal judgments might be due to their being more difficult, Bressler (14) attempted to equalize the difficulty of the three categories by demanding an identification of *each* stimulus, instead of only the middle stimulus in the range as in the usual procedure. For this purpose he performed an experiment using eleven comparison weights and requiring the Ss to identify each stimulus as 84 gm., 108 gm., etc. All identifications greater than 100 (standard) were grouped as 'heavier' judgments, all less than 100 as 'lighter', and all 100 gm. identifications as 'equal', permitting calculation by the usual methods. Comparing the results by this method with those of a control conducted under the ordinary procedure, Bressler found a higher precision for the distributions of heavier and lighter judgments in the new method,[4] although he was unable to present positive evidence that the method equalized the difficulty of judgment in the various categories.

A different line of attack upon the problem of the 'equal' category was opened by Kellogg (85), with the assumptions that if the equal category is allowable it must be conditioned by factors similar to those of the extreme categories, and that similar sets of factors should result in similar reaction times. No decisive experiment having been performed upon this problem, Kellogg compared the reaction times for a two-category experiment with those of a three-category experiment, the judgments being made on visual brightness. He reported significantly longer times for the three-category plan than for the two-category plan. 'Equal' judgment-times were significantly longer than those for difference judgments, and correct judgments were quicker than incorrect ones. When time curves were plotted in a manner similar to the psychometric functions they were

[4] There appears to be an error in the author's statement of his second conclusion, page 43, where "standard deviations" should evidently be read for "precision measures".

found to resemble irregular inversions of these functions. Kellogg stated that "within any given category of judgment, the more the stimuli deviate from the ideally perfect condition typified by the category name, the longer the judgment-times tend to become" (p. 81). This means, of course, that equal judgments require more time when the stimulus difference is large than when it is small, which Kellogg interpreted as indicating the presence of a disturbing factor in such judgments; he concluded that equal judgments were slower than difference judgments even when no element of doubt was present.

Fernberger and Irwin (**46**) studied the same problem by the method of single stimuli with the use of lifted weights. Their results failed to confirm Kellogg's finding of a considerable difference in time for 'equal' as against 'difference' judgments. The conclusion was drawn that insofar as the reaction time criterion was valid, the category 'equal' was entirely permissible with the method of single stimuli. In a further analysis of the results Irwin and Fernberger (**74**) completely confirmed Kellogg's finding of the time curves as inversions of the psychometric functions. The work of Carlson, Driver and Preston (**17**) and Fernberger, Glass, Hoffman and Willig (**45**) gave evidence that the discrepancy in the findings on time-differences was due to the subject's attitude toward equal judgments. An attitude unfavorable to these judgments was associated with a lengthening of the reaction-time which was not present with a favorable attitude. Because of the peculiarities of his method the significance of Hamilton's (**65**) conclusion that the time of judgment had little relation to its accuracy is doubtful.

We have seen above that Culler asserted the presence of finer sensitivity with the inclusion of equal judgments than with their exclusion. Kellogg (**84**) stressed the fact that Culler's opinion was based upon statistical treatment rather than upon results of a direct experimental attack. He therefore performed an experiment using two- and three-category plans with the same Ss under carefully controlled conditions, his instructions not emphasizing any one category but rather permitting the Ss to distribute their 'doubtful' cases as they thought best. In the 'equal' category, however, the Ss were to include only cases of absolute equality. When the σ's of the distributions of the different categories were calculated according to Spearman's process, those for the three-category plan were found slightly smaller than the corresponding σ's of the two-category plan. The differences were not reliable in individual cases. A well marked negative correlation between the ranking for the Ss as to the number

TABLE I

SPECIAL INFLUENCES UPON THE JUDGMENT

Influence	Purported Effects, etc.	Evidence	Reference
Practice	If 'general' causes tendency toward agreement with $\Phi(\gamma)$ hypothesis; if 'specific' no such tendency.	Logical and experimental.	Culler (26)
	Decrease in number of equal judgments of all degrees of assurance.	Lifted weights experiment.	Fernberger (43)
	Effect of practice in visual discrimination of length of lines very slight.	Experiment with modification of Galton bar.	Hamilton (65)
	Slight decrease in upper limen for tone.	Experiment with Koenig cylinders.	Humes (72)
	Difficulty of eliminating effects of practice in psychophysics.	Visual and auditory experiments.	Kellogg (83)
	Effects may be due to decrease in 'discriminal dispersions'.	Theoretical.	Thurstone (136)
	When practice is apparent within a series the method of constant stimuli cannot be used.	Theoretical.	Urban (162)
Attitude	Affects judgment reaction time.	Theory to explain conflict among experimental results.	Carlson *et al.* (17); Fernberger *et al.* (45)
	Stimulus attitude leads to greater sensitivity.	Results of various experimenters.	Culler (25)

	Attitude can be maintained over period of five years.	One case.	Fernberger (40)
	Difference in attitude may be due to temperamental factors.	Theory.	Fernberger (42)
	Scatter of equal judgments depends upon attitude.	Theory.	Fernberger (42)
	Attitude toward 'whole' favors low DL for 'whole'.	Experimental.	Irwin (73)
	Shifts in attitude more important in average error than constant stimuli.	Visual and auditory experiments.	Kellogg (83)
	Difference in attitude between methods of average error and constant stimuli.	Same as above.	Kellogg (83)
	Constancy of attitude very uncertain in psychophysical experiment.	Lifted weights experiment.	Rudisill (128)
Instructions	'Formal' analysis of instructions possible.	Lifted weights.	Fernberger (42)
	Instructions can be framed to include under middle category whatever is desired.	Same as above.	Fernberger (42)
Previous or interpolated stimuli or judgments, and temporal configuration	Existence of tendency to avoid repetition of judgments in given category.	Equal weights experiment.	Arons and Irwin (2)
	First weight of pair always heavy in comparison with preceding pairs.	Personal experience.	Culler (25)
	Interpolated stimulus increases DL and disturbs psychometric functions.	Lifted weight experiment.	Guilford and Park (63)
	Previous stimulus affects probabilities of judgments on succeeding stimuli.	Lifted weight experiment.	Turner (159)

TABLE I—*Continued*

Influence	Purported Effects, etc.	Evidence	Reference
	Various temporal configurations of stimuli result in tendencies of 'assimilation' or 'opposition'.	Tone intensities experiment.	Onoshima (109)
Subliminal stimulus	Subliminal stimulus effective in judgment.	Experiment with Müller-Lyer illusion.	Bressler (13)
Fatigue	In method of average error fatigue made settings close to previous settings.	Experimental observation.	Kellogg (83)
	Review of work on fatigue as affecting the threshold.	Historical.	Spencer and Cohen (133)
	Positive correlation between 'freshness' and visual limen. (But *cf.* Geldard, 54.)	Visual experiment.	Spencer and Cohen (133)
Central	Effect of central factors very important in judgment as compared with that of peripheral factors ('transformation effect').	Experimental results.	Fischel (47)
Gradual stimulation	Effect of gradual stimulation (by weights) less than that of sudden stimulation.	Experimental results.	Walter (170)
Incentives	Result in greatly increased accuracy of discrimination.	Visual discrimination of length of lines.	Hamilton (65)

of equal judgments they gave and the averaged σ's of their distributions for the extreme categories was noted. This had already been observed by Culler, as mentioned above, who explained it on the basis of the mathematical relationship between the number of equal judgments and the slope of the psychometric functions of the extreme categories. The reliabilities of the two plans were not significantly different. The author's most important conclusion was that "within a reasonable PE the precision measures [of the two plans] can be regarded as practically equal" (p. 42).

The conclusion to be drawn from this discussion may be put briefly: Although there are many reasons for mistrusting the equal or intermediate category of judgment in the method of constant stimuli, no experimental results have conclusively demonstrated the necessity for its exclusion. Hence, whenever adequate control of attitude and other subjective factors is possible it may be used. Unfortunately, however, the word 'adequate' is easier to write than to define.

Special Influences upon the Judgment. For the sake of conciseness the results of investigations dealing with certain influences upon the psychophysical judgment have been gathered into Table I. Some of these factors have been considered of great theoretical importance by those who studied them, so that thorough justice has not always been done them by this brief mention. It may be noted that not all of the influences tabulated are equally well supported by experimental evidence.

Statistical Considerations. Of the various psychophysical methods that of constant stimuli has been treated most thoroughly from the statistical point of view. Although much attention has been paid to the facilitation of the practice of this method, the calculation of the constants of the psychometric functions by the Müller-Urban process is so laborious, even with the tables of Urban and Rich, that several attempts have been made to justify the use of simpler methods. Newhall (**106**) suggested that linear interpolation could be used to determine the limen and its PE, with a great economy of the computer's time and labor. Calculating limens for 100 distributions, he found that the correlations between interpolated limens and limens obtained by the constant process ranged from .960 to .987 according to whether the interpolation procedure did or did not assume normality. On lifted weight results of his own, where five comparison stimuli were used, Newhall (**107**) showed that the interpolated limens differed less from the constant process limens than the latter differed from each other, the differences in any case being insignificant.

Similar results were obtained when only three comparison stimuli were used; no important differences between the procedures with five and three stimuli were found, but the data were not too extensive. No data were presented on the correlation between the PE's of the methods compared. Newhall's general conclusion was that the interpolated limens were comparable to those calculated by the constant process.

A further step was taken by Linder (97), who compared the interpolation and constant processes by a purely statistical method which limitations of space unfortunately prevent us from describing. Using as a criterion the scatter around a theoretical value yielded by his method, he found that linear interpolation between two points immediately around the median, and similar interpolation on the integral of the normal curve, were significantly inferior to both Müller's and Urban's methods. Here, it must be observed, the interpolation methods used but a fraction of the actual data; but when the same number of observations were concentrated at two points the interpolation methods were somewhat better than the others, with linear interpolations slightly superior to normal curve interpolation. Linder's results were, of course, entirely 'formal' in Urban's sense; he intentionally avoided raising the issue of *psychological* differences.

In practice, the experimenter who must choose among methods of calculation cannot apply Newhall's and Linder's conclusions immediately to his own problem. As Brown and Thomson (16) have pointed out, he must consider the relative accuracy of his experiment in order to avoid the common error of applying too refined a method to data too little refined. If he decides upon the constant process he must assure himself that the data are neither heterogeneous nor significantly skew in distribution. If he wants to employ Spearman's arithmetic mean he must see to it that by the use of appropriate stimuli the relative frequencies at the extremes approach at least very nearly the values zero and one. On the other hand, any economy of time and labor naturally is unjustified if in any way it results in misrepresentation of the data. Unquestionably, however, the work of Newhall and Linder has improved the status of the interpolation methods with their attendant savings.

An interesting discussion of the usefulness of the normal law in psychological investigation in general and psychometric work in particular has been carried on by several writers. Brown and Thomson (16) retained in this edition the statements that "... we are led to believe that the difficulties of psychophysical experiment are such that homogeneity in the data is rare. For such data refine-

ments of mathematical calculation are out of place" (p. 94), and that "very few collections of psychophysical data are worth the accuracy of the constant process, which is however undoubtedly the best theoretically for symmetrical distributions" (p. 75). Again, Boring[5] had criticized the attempt to discover a generalized psychometric function, that is, one generally valid without reference to the particular stimuli used or to their scale of measurement; and he denied that practice had been shown to make data tend toward normality. He emphasized the point that the generality of application of the normal curve was limited by non-linear relations among measurements of the same object, *e.g.*, that weight might be supposed to be roughly proportional to volume and hence to the cube of height.

We shall discuss first the matter of how one should determine the approximation of psychophysical data to normality or to the phi-gamma hypothesis. Hoisington[6] and Fernberger[7] had used as their criterion the sum of the squared unweighted deviations of the experimental from the hypothetical relative frequencies. This measure was criticized by Rich (121) on apparently adequate grounds. Rich and all later writers prefer Thomson's criterion in which the squares of the deviations are weighted. Since he was able to show that in some cases they lead to opposite conclusions, Rich demonstrated that a serious choice between the two criteria was necessary. At present there seems no doubt that Thomson's criterion is correct if, as does Culler (26), one defines the best-fitting curve as the most probable one. It must be kept in mind, of course, that 'most probable' refers only to one curve among those actually tried for fit; thus, if one is fitting the normal curve the 'most probable curve' is simply the most probable *normal* curve for the given data.

Boring's problem of the unit in psychophysical measurements was considered by Culler (27), according to whom it becomes unnecessary to worry about the matter when we realize that our interest is in units of energy or some linear function of energy. We need only measure physically the energy of the stimulus; and if this cannot be done directly, Culler believes that it can be done indirectly by means of the organism's responses. The reviewer agrees with this argu-

[5] Boring, E. G., Is There a Generalized Psychometric Function? *Amer. J. Psychol.*, 1924, **35**, 75–78.

[6] Hoisington, L. B., An Example of the Fractionation of Data from the Method of Constant Stimuli for the Two-Point Limen. *Amer. J. Psychol.*, 1917, **28**, 588–596.

[7] Fernberger, S. W., The Approximation of Actual Data to the Phi-Gamma Hypothesis. *Amer. J. Psychol.*, 1923, **34**, 498–500.

ment as it applies specifically to certain of Boring's illustrations; but he is unable to absolve from the suspicion of circularity that part of it concerned with measuring energy by means of response, since it seems impossible to interpret the behavior without previous knowledge of the stimulus energy. Aside from this, Culler showed in the same paper that contrary to expectation each of two measurements related to each other theoretically in a non-linear way may still be

TABLE IIA

APPLICATION OF THE χ^2 CRITERION TO CULLER'S DISTRIBUTION OF VALUES OF PEARSON'S P FOR 300 PSYCHOMETRIC DISTRIBUTIONS

Pearson's P	Number Observed	Number Expected	Difference	Difference ^{2} ÷ Number Expected
1.00–0.75	116	75	41	22.41
0.74–0.50	57	75	—18	4.32
0.49–0.25	62	75	—13	2.25
0.24–0.10	26	45	—19	8.02
0.09–0.05	12	15	—3	.60
Below 0.05	27	15	12	9.60
Totals	300	300		47.20

For $\chi^2=47.20$ and $n=5$, $P<.01$ (considerably).

TABLE IIB

APPLICATION OF THE χ^2 CRITERION TO CULLER'S DISTRIBUTION OF VALUES OF PEARSON'S P FOR 300 PSYCHOMETRIC DISTRIBUTIONS

With Data Grouped in Two Classes

(1.00–0.75 and Below 0.05, Taken Together, and 0.74–0.05)

Pearson's P	Number Observed	Number Expected	Difference	Difference ^{2} ÷ Number Expected
1.00–0.75 and Below 0.05	143	90	53	31.21
0.74–0.05	157	210	—53	13.38
Totals	300	300		44.59

For $\chi^2=44.59$ and $n=1$, $P<.01$ (considerably).

distributed normally in practice. Thus, though the weight and breadth of eggs, to use one of his examples, are in theory related as x to x^3, both weight and breadth when actually measured give good fits to normal curves. The same is true for the length and weight of infants at birth. Culler explained this as due to the departures of natural objects from ideal relations.

In further defense of the usefulness of the normal law, Culler (27) selected 300 psychometric distributions from the work of well-known experimenters and classified them on the basis of Pearson P-values for goodness of fit of the normal curve (see our Tables II A, B, C,

and D). He concluded that under good conditions with a careful observer *" we are more likely to find a normal sequence in his reports than when drawing balls from a 50–50 urn "* (p. 290; italics Culler's). Since the whole question is of considerable importance, it may be well to examine it somewhat further. In the first place, we have applied the χ^2 criterion to the values of the above table with the resulting discovery that the deviation from expectation is significant

TABLE IIc

APPLICATION OF THE χ^2 CRITERION TO CULLER'S DISTRIBUTION OF VALUES OF PEARSON'S P FOR 300 PSYCHOMETRIC DISTRIBUTIONS

With Data Grouped in Two Classes

(1.00–0.75 and Below 0.75)

Pearson's P	Number Observed	Number Expected	Difference	Difference ^{2} ÷ Number Expected
1.00–0.75	116	75	41	22.41
Below 0.75	184	225	—41	7.47
Totals	300	300		29.88

For $\chi^2=29.88$ and $n=1$, $P<.01$ (considerably).

TABLE IID

APPLICATION OF THE χ^2 CRITERION TO CULLER'S DISTRIBUTION OF VALUES OF PEARSON'S P FOR 300 PSYCHOMETRIC DISTRIBUTIONS

With Data Grouped in Two Classes

(1.00–0.05 and Below 0.05)

Pearson's P	Number Observed	Number Expected	Difference	Difference ^{2} ÷ Number Expected
1.00–0.05	273	285	—12	.51
Below 0.05	27	15	12	9.60
Totals	300	300		10.11

For $\chi^2=10.11$ and $n=1$, $P<.01$ (considerably).

(First three columns in Tables IIA, IIB, IIc, and IID from Culler, **27**, p. 289. On the statistical device used by the reviewer see Yule, G. U., *An Introduction to the Theory of Statistics.* London, 1932, 377–378.)

for: (1) all six classes taken together, and (2) the middle four classes together. Further, by the same criterion, there is a significant excess beyond expectation: (3) for the first class (P's above .75), and (4) for the last class (P's below .05). In each of these cases the deviation from expectation is so great as to make it highly improbable that it is due to fluctuations of random sampling. Thus, not only do we not have a random sampling of P-values, but there are also too many both of very low and very high values, as stated by Culler. Probably no one will be astonished by the large number of very low

values, but the excess of values above .75 is hard to account for. Inspection of the results of the separate experiments as reported by Culler shows that this excess is not due to any one type of material; it is found both with temperatures and weights, though not clearly seen in the one auditory experiment. Nor is it due to any small group of observers; most of the subjects were contributory to it.

In this connection it may be well to consider two influences which have been shown to be important with respect to expectation of normality. First is the observation made by Thurstone (**145**) which, surprisingly enough, appears not to have been made before, to the effect that in cases where the absolute limen is an increasing function of the stimulus, the psychometric functions for the two extreme categories of judgment must be positively skewed, and hence cannot be fitted exactly by any symmetrical curve. If Weber's law is valid, a Φ (log γ) hypothesis will hold, but Thurstone's argument does not depend upon the assumption of Weber's law. Fritz (**51**) gave two examples of the application of the argument to experimental results, somewhat unnecessarily, it appears, since Thurstone's contention is based upon logical and mathematical reasoning and can be neither supported nor controverted by experimental evidence. Thurstone limited his discussion to the case of experiments from which the intermediate category of judgment was excluded, but it seems safe to expect some asymmetry also where this category is included.

The second influence upon the normality of the psychometric function is that of practice. The question as to whether the $\Phi(\gamma)$ hypothesis is more nearly realized with increased practice is undoubtedly very important for the hypothesis itself. Culler (**26**) maintained that we must distinguish between (a) general practice in observation and (b) specific practice at a given task. Given (a), although h may increase, there is no reason to expect closer approach to normality; but during the acquirement of (a) normality should be approached. He showed further that goodness of fit (P) is related to precision (h); when h is high, P must also be high, but when h is low, P may be either high or low. From a careful examination of the experimental evidence Culler drew the conclusion that the above statements were substantiated; in general, the $\Phi(\gamma)$ hypothesis was the limiting form approached by sensory data with practice.

We now return to Culler's distribution of P-values. The influence discussed by Thurstone would lead us to expect departures from normality of the individual cases, though relatively slight departures because of the high precision of the results in question. This might be one factor in the large number of very low P's. As for practice,

the eleven observers whose results are here discussed were known to be at different stages. Those who were at an early stage would give results which, again, would increase the number of low P values. Here, however, there is a bias of some kind which brings even more high values than can be accounted for by random sampling from a strictly normal universe. From Culler's discussion it appears that he attributes this to the high degree of training of some of the observers. However, it is hard to see how even perfected training (using the word 'training' in the ordinary sense) could do more than raise the goodness of fit to that expected from random sampling. The reviewer finds himself unable to explain this state of affairs but he feels that, at any rate, Culler's (27) picture of the observer as an energy system of balanced forces of which normality is the expectation is surely not validated by these results—if such an image is the true one, it remains to be shown wherein the bias lies.

Several papers have dealt with the probable error of the limen of the method of constant stimuli. Urban (160) admitted that although the method of constant stimuli was the "best for the experimenter" it was the "least accurate of all" methods, though he was puzzled as to the reason for this. Culler (23, 25) maintained that Urban's derivation of the formula for PE_L was "fatally defective", since it made Urban's own results highly improbable and was altered as to the values it gave when the origin was shifted. The explanation of these difficulties according to Culler was Urban's assumption that c and h were uncorrelated when, as Culler showed, c and h were experimentally correlated. The PE_L suggested by Boring [8] was criticized by Culler on the basis of its being a PE of an average while the limen is defined as a median. Lufkin (98) countered by showing that if c and h are not independent, the method of least squares breaks down and the $\Phi(\gamma)$ hypothesis must be discarded. To this Culler (24) replied that he had not maintained that these values were correlated in the mathematical sense and hence the $\Phi(\gamma)$ hypothesis was not involved. He went on to show that, according to conditions, c and h can be correlated from 0.00 to 1.00. The formula suggested by Culler is: $PE_L = .84535(1/h\sqrt{2n})$. This comes to about 4/3 of Thomson's PE_L, the difference being explained by Culler (25) as chiefly due to Thomson's use of Urban's weights which, as he showed in the same paper, fail in some cases to be close enough approximations for desirable accuracy. In a further statement of

[8] Boring, E. G., The Number of Observations on Which a Limen May Be Based. *Amer. J. Psychol.*, 1916, 27, 315–319.

the problem Culler (**28**) gave a generalized formula in which the $PE_L = .84535(1/h\sqrt{2\Sigma nPp})$, in which P represents the Urban weights and p the judgment percentages, and ΣPp is not necessarily equal to unity as in the former more limited formula. Testing this formula by applying it to Urban's data, he found the actual dispersion greater than that predicted for relatively untrained subjects, but approximately equal to it for well-trained subjects. As the final outcome of this discussion it appears, in the reviewer's opinion, that Culler's position throughout the discussion is well taken.

Lufkin (**99**) has the distinction of being the first to compare the fit of normal, Pearsonian, and Gram-Charlier curves to Urban's familiar data. There appeared to be little choice as to goodness of fit, what advantage there was being in favor of the Gram-Charlier, followed by the Pearsonian and normal curves. Lufkin expressed the belief that, had the higher moments been adjusted by the method of least squares, the Gram-Charlier curves would have given the best fit in all cases. Thomson (**135**) criticized Lufkin for failing to make explicit his assumptions as to the 'tails' of the distribution. He reiterated the statements on the importance of this problem made in Brown and Thomson (**16**), and set forth the various assumptions that could be made with tests of their sufficiency.

The importance in psychometric theory of Xi, the point of subjective equality, was stressed by Culler (**25**). He held that the limen should be measured from this point rather than from the standard stimulus. Furthermore, he recommended the use of the PE of the psychometric function for the limen, instead of the median of this function as employed by Urban. A similar suggestion was made on different grounds by Wever and Zener (**176**). The advantages claimed by Culler for the PE were its greater precision—more than 2½ times that of the median—and its simplicity and convenience. It occurs to the writer, however, that its meaning as a limen is less clear than that of the median, and that its use would tend to raise a vexatious problem in the constant method similar to that which has been so troublesome in the method of average error, namely, how a measure of scatter is related to the limen or just noticeable difference.

In the constant method as commonly practiced, comparison stimuli are compared with a single standard stimulus and the judgment probabilities are considered as functions of a single variable. By mathematical development of the psychometric functions, Urban (**164, 166, 167**) generalized the method so as to consider the judgments as functions of two variables, the standard and comparison stimuli. This treatment he considered necessary in using the constant

method for such a problem as Weber's law. Applying this method to Thurstone's results from the method of equal-appearing intervals, Urban concluded that since the results obtained with the use of the generalized method agreed well with those obtained without it, the 'mental manifold' or series of sensation-intervals, is measurable. If the reviewer understands Urban's argument, the question as to the measurability of sensation-scales is here made to depend upon the $\Phi(\gamma)$ hypothesis. If this is so, further discussion to demonstrate the validity of this criterion is to be hoped for, particularly since it would seem undesirable to make such an important question directly depend upon so specific an experimental hypothesis (or even, as some might say, upon any experimental hypothesis at all).

General and Theoretical Considerations. In a symposium at the Eighth International Congress Boring (**9**), Myers (**104**), and Werner (**173**) discussed the nature of sensation-intensity. Later, Johnson (**76**), raising a fundamental theoretical issue, declared that Fechner's attempt at the measurement of sensation was founded upon equivocation, because the formula $\gamma = k \log \frac{\beta}{b}$ defined merely a function of the ratio of two stimulus intensities. Replying to Johnson, Boring (**10**) maintained that the formula represented a relation between two empirical quantities, in that the just noticeable difference of sensation is observable and "may even be observable as just noticeable" (p. 444). He asserted that the j.n.d. "is not a difference of stimuli" but "a difference in sensation" (p. 443). Returning to the attack, Johnson (**77**) upheld his belief that a 'sensation' could not be measured as independent of the measurement formula, and that consequently the formula was *definitional.* He denied the direct observability of the j.n.d., holding that it was simply a statistical concept derived from a census of classified responses, and in a further discussion (**78**) he pointed out certain properties of γ (sensation-intensity) in the *Massformel* which could not hold of 'introspectional' sensations. To the interested spectator this conflict of opinion may well present itself as a symptom of the battle to the death being waged at the present time on the question of 'objectivity' in psychology. One may well suspect that there is present here a critical conflict on an epistemological, rather than a psychological, problem.

No doubt the same observer is anxiously awaiting the event of general agreement, since the possibility that Fechner's 'law' is definitional appears still to be disregarded completely by experimenters accumulating data to attack or defend its empirical validity,

or writing theoretical articles in attempts at explanation of it. The significance, for example, of all such work as that of Cobb (**20**), Zinner (**180, 181**), Pauli and Wenzl (**110**), Hoagland (**69**), Peddie (**111**), Houstoun and Shearer (**70**), Guilford (**61**), Bénézé (**7**), and others must remain very doubtful so long as the final solution is lacking. If one holds with Johnson and with Singer (see an important paper by Lienau, **96**) that the law is definitional, there is, of course, no question of 'verification', experiment being left merely the task (weighty enough in itself) of determining whether anything exists which fulfils the definition.

A clearer understanding of the relation between Weber's and Fechner's laws has recently come into being, as exemplified by Boring's (**11**), Thurstone's (**136**), and Urban's (**165, 167**) treatment of the subject. The latter two make especially plain their belief that the two laws are mutually independent in the sense that data might conceivably be found in which either one was verified and the other not; at any rate, the two formulas are not now so often treated as identical in meaning (*cf.* Fechner's own discussion [9]). It is strongly advisable that Boring's suggestion as to terminology be generally accepted, namely, that the formula $\frac{\delta R}{R}$=const. be known as 'Weber's' law and the formula $S=k \log R$ as 'Fechner's' law.

The nature of the threshold has been investigated by several authors. Using a weak visual stimulus near the stimulus limen Guilford (**57**) found that the per cent of time during which it was visible was a function of the intensity of the stimulus and that this function was well fitted by the $\Phi(\gamma)$ curve. He was able to write: "Thus it appears that when we have found the factors which favor fluctuation we shall also have found the factors which control the limen" (p. 554). The point of 50 per cent visibility, when allowance was made for local adaptation, was concluded to be near the limen as determined by the usual procedure.

The work of Anderson (**1**), Spencer (**130, 131**), and Spencer and Cohen (**132, 133**) which led toward the conclusion that the visual threshold was to be explained in terms of Heymans' law of inhibition, was criticized by Geldard (**54**) on the grounds that the results, when properly corrected for pupil diameter, led to a quite different conclusion.

The influence of the development of the all-or-none law upon the concept of the threshold has been great. It may be seen at its

[9] Fechner, G. T., *Elemente der Psychophysik.* Dritte unveränd. Aufl. Leipzig, 1907, I, 57, and II, 2.

clearest, perhaps, in the work of Hecht (**66, 67**) (on which *cf.* Graham and Goldman, **56**), and Hoagland (**69**), where sensation-intensity was taken to be associated with the number of receptors active, and confirmation of this theory was sought in the numerical coincidence between the number of sensation-steps between the stimulus and terminal limens and the number of receptors available for these steps as calculated on independent grounds. Hoagland carried the argument to an almost alarming extreme in explaining Fechner's law and curves of learning. A nice question is involved here as to the validity—apparently implied by Hecht and Hoagland—of treating the threshold as an absolute, rather than a conventional, value. The reviewer hopes to discuss this problem more fully in a later paper.

The existence of the threshold was explained by Carr (**18**) on the basis of the non-equivalence of the units of measurement of stimulus and sensation. We can measure the stimulus very exactly, he said, but "the differences in sensory effect are directly measured by an *unaided* sense organ" (p. 317). We may ask here whether it is not relevant to say that such sensory differences are not measured by a sense organ at all, but rather by a statistical classification of the organism's responses. Is it not necessary to define 'sensory effect' in such statistical terms? Certain relationships to the issue between Boring and Johnson, discussed above, are obvious.

An important contribution to the analysis of constant errors was made by Culler (**25**) with his exposition of a movement-error, overlooked in the classical treatment of time- and space-errors, but defined by the same method. In the same paper Culler presented a criticism and theory of the concept of absolute impression.

The findings of Ponzo and Angyal (**115**) which led them to suggest that there are two distinct sensations of weight, 'light' and 'heavy', related to one another somewhat as cold to warm, are of possible interest in connection with absolute impression and the time-error as found in lifted weight experiments.

The tremendous influence of the Gestalt psychology has involved it in almost all of the significant contributions to the theory of the comparison judgment during this period. Chiefly instrumental was the work of Köhler (**91**) who, making his point of departure an experiment of Borak's [10], established so firmly the importance of the theory of the time-error in relation to that of the judgment that

[10] Borak, J., Ueber die Empfindlichkeit für Gewichtsunterschiede bei abnehmender Reizstärke. *Psychol. Forsch.*, 1922, 1, 374–389. Since it was omitted from Rich's review we include Köhler's paper in our bibliography although it appeared in 1923.

the two theories have since become inseparable. We need not discuss this work at any length since it has been ably covered by Needham (105) in a thorough historical review. Among the many investigations inspired at least in part by Köhler's work may be mentioned those of Lauenstein (95), Wada (169), Pratt (116, 117), Guilford and Park (63), Turner (159), and Woodrow (178). For a complete bibliography and discussion see the work of Needham cited above.

New Applications, etc. The field of usefulness of psychophysical methods has been greatly enlarged by new applications. Rich (122) and Dimmick and McMichael (35), for example, argued for the status of pure gray as an unique experience on the basis of their finding that limens of black from gray and white from gray could be obtained. Katz and Noldt (79) upheld the theory of the independent existence of a sense of vibration, their evidence being the very low limen for the perception of vibratory movement by the fingertips (about 0.001 mm.) which they maintained was too small a movement to be perceived by the sense of touch as ordinarily understood.

Other applications of psychophysical methods to various types of problems were those of Guilford and Dallenbach (62), determination of memory span by the method of constant stimuli; White and Landis (177), limens for the perception of differences in a silhouette face; Hallett (64) and Irwin (73), in relation to the Gestalt theory of perception; Gengerelli (55), measurement of eidetic imagery by the method of constant stimuli; Fernberger (39), individual differences; Tiffin (154), study of the vibrato; Travis (157), diagnosis of character types; and Wendt (172) and Barker (5), industrial applications.

An analysis of psychophysical measurement capable of the broadest application was made by Thurstone (136). Taking an entirely original approach Thurstone developed a formula, the eminently practical value of which (contrasting so greatly with 'conventional psychophysics') lay in its applicability to stimuli of a high degree of complexity or resistance to physical measurement. While Thurstone wrote of his final result as 'the law of comparative judgment' it appears that it is definitional in nature by the criterion applied by Johnson (cited above) to Fechner's 'law'. Many psychophysicists will no doubt feel the force of the incentive leading to Thurstone's analysis, as stated by himself: "Several years ago when I was teaching conventional psychophysics, it seemed to me that psychophysics was really a very dull subject in spite of the fact that it did offer the satisfaction of clean and quantitative logic. This type

of satisfaction is rare in psychological investigation and consequently psychophysics has stood out as a very dignified topic in psychology in spite of the fact that its intrinsic subject matter has been, on the whole, rather trivial" (151, p. 249). No complete theoretical evaluation of the psychophysical foundations of this work has been found by the reviewer, although it would be surprising if psychologists had no objections to raise to such far-flung extensions of fundamental concepts. The one psychologist, Guilford, who has made considerable use of Thurstone's ideas, raised the question as to whether the results obtained so laboriously by applying Thurstone's definition and methods of measurement were commensurate with the labor involved. Testing a quicker method for determining scale values, Guilford (59) found it to be superior to Thurstone's in ease of application and in validity. He found also that the correlation between Thurstone's scale values and the mere numbers of preferences was very high—about .99. Workers in the social sciences have been quick to take note of Thurstone's methods; see, for example, the comments of Rice (119), replied to by Thurstone (148). Brief reviews are to be found in Bain (4) and Droba (36); the latter contains additional references, the present bibliography listing mostly the more purely theoretical articles. The relevant items on our list are: Thurstone (137, 138, 139, 140, 141, 142, 143, 144, 147, 149, 150, 151, 152), Thurstone and Chave (153), Peterson and Thurstone (112), Guilford (58, 59, 60), Farnsworth (37), Farnsworth and Chichizola (38), Hevner (68), Folgmann (49), and a correction by Babb (3).

BIBLIOGRAPHY

1. Anderson, J. E., On Inhibition of the Differential Threshold in Accordance with Heymans' Law. *Amer. J. Psychol.*, 1925, 36, 461–463.
2. Arons, L., and Irwin, F. W., Equal Weights and Psychophysical Judgments. *J. Exper. Psychol.*, 1932, 15, 733–751.
3. Babb, W. W. M., Correction. *J. Exper. Psychol.*, 1933, 16, 611.
4. Bain, R., Theory and Measurement of Attitudes and Opinions. *Psychol. Bull.*, 1930, 27, 357–379.
5. Barker, S. G., The Fechner-Weber Law in Wool Sorting. *Nature*, 1930, 126, 646.
6. Beebe-Center, J. G., *The Psychology of Pleasantness and Unpleasantness.* New York: Van Nostrand, 1932, 12–37, 196–199.
7. Bénézé, G., Notes sur la loi de Fechner. *Rev. phil.*, 1929, 108, 429–432.
8. Blumenfeld, W., Urteil und Beurteilung. *Arch. f. d. ges. Psychol.*, Dritter Ergänzungsband, 1931. Pp. x+623.
9. Boring, E. G., The Intensity of Sensation. Proc. VIII Internat. Congr. Psychol. Groningen: P. Noordhoff, 1927, 71–78.

10. BORING, E. G., Did Fechner Measure Sensation? *Psychol. Rev.*, 1928, **35**, 443–445.
11. BORING, E. G., *A History of Experimental Psychology*. New York and London: Century Co., 1929. Pp. xvi+699.
12. BORING, E. G., The Two-Point Limen and the Error of Localization. *Amer. J. Psychol.*, 1930, **42**, 446–449.
13. BRESSLER, J., Illusion in the Case of Subliminal Visual Stimulation. *J. Gener. Psychol.*, 1931, **5**, 244–251.
14. BRESSLER, J., Judgment in Absolute Units as a Psychophysical Method. *Arch. of Psychol.*, 1933, No. 152. Pp. 68.
15. BROWN, M. S., and STRATTON, G. M., The Spatial Threshold of Touch in Blind and in Seeing Children. *J. Exper. Psychol.*, 1925, **8**, 434–442.
16. BROWN, W., and THOMSON, G. H., *The Essentials of Mental Measurement*. (3d Ed.) Cambridge: Cambridge Univ. Press, 1925. Pp. x+224.
17. CARLSON, W. R., DRIVER, R. C., and PRESTON, M. G., Judgment Times for the Method of Constant Stimuli. *J. Exper. Psychol.*, 1934, **17**, 113–118.
18. CARR, H., An Interpretation of the Weber-Fechner Law. *Psychol. Rev.*, 1927, **34**, 313–319.
19. CARTZDAFNER, J. P., An Experimental Comparison of Psychophysical Methods. *Psychol. Bull.*, 1932, **29**, 570–571.
20. COBB, P. W., Weber's Law and the Fechnerian Muddle. *Psychol. Rev.*, 1932, **39**, 533–551.
21. CULLER, E., On Thermal Sensitivity and the Nature of Sensory Adaptation. *Brit. J. Psychol.*, 1926, **16**, 193–198.
22. CULLER, E., Thermal Discrimination and Weber's Law. *Arch. of Psychol.*, 1926, No. 81. Pp. 134.
23. CULLER, E., The Accuracy of the Method of Constant Stimuli: a Reply to Dr. Urban. *Amer. J. Psychol.*, 1927, **38**, 307–312.
24. CULLER, E., The Accuracy of Müller-Urban's Method of Constant Stimuli. *Amer. J. Psychol.*, 1928, **40**, 342–345.
25. CULLER, E., Studies in Psychometric Theory. I–X. *Psychol. Monog.*, 1926, **35**, No. 2, 56–137.
26. CULLER, E., Studies in Psychometric Theory. XI and XII. *J. Exper. Psychol.*, 1926, **9**, 169–194.
27. CULLER, E., Studies in Psychometric Theory. XIII. *J. Exper. Psychol.*, 1926, **9**, 271–298.
28. CULLER, E., Studies in Psychometric Theory. XIV. *J. Exper. Psychol.*, 1927, **10**, 463–477.
29. DALLENBACH, K. M., Attention. *Psychol. Bull.*, 1926, **23**, 1–18.
30. DALLENBACH, K. M., Attention. *Psychol. Bull.*, 1928, **25**, 493–512.
31. DALLENBACH, K. M., Attention. *Psychol. Bull.*, 1930, **27**, 497–513.
32. DALLENBACH, K. M., Tables for the Constant Methods. *Amer. J. Psychol.*, 1933, **45**, 149.
33. DAVIES, E., and BABB, W. W. M., Concerning the Absolute Method in Psychophysics. *Amer. J. Psychol.*, 1932, **44**, 367–369.
34. DIMMICK, F. L., and KARL, J. C., The Effect of the Exposure Time upon the R. L. of Visible Motion. *J. Exper. Psychol.*, 1930, **13**, 365–369.

35. DIMMICK, F. L., and MCMICHAEL, G., The Psychophysical Determination of the Limits of Pure Gray. *Amer. J. Psychol.*, 1933, **45**, 313–314.
36. DROBA, D. D., Methods for Measuring Attitudes. *Psychol. Bull.*, 1932, **29**, 309–323.
37. FARNSWORTH, P. R., Preferences for Rectangles. *J. Gener. Psychol.*, 1932, **7**, 479–481.
38. FARNSWORTH, P. R., and CHICHIZOLA, T. L., Color Preferences in Terms of Sigma Units. *Amer. J. Psychol.*, 1931, **43**, 631.
39. FERNBERGER, S. W., An Unconsidered Source of Material for the Problem of Individual Differences. *Amer. J. Psychol.*, 1930, **42**, 646–647.
40. FERNBERGER, S. W., Constancy of Attitude Maintained Over a Long Period of Time. *Amer. J. Psychol.*, 1930, **42**, 317–318.
41. FERNBERGER, S. W., The Use of Equality Judgments in Psychophysical Procedures. *Psychol. Rev.*, 1930, **37**, 107–112.
42. FERNBERGER, S. W., Instructions and the Psychophysical Limen. *Amer. J. Psychol.*, 1931, **43**, 361–376.
43. FERNBERGER, S. W., Die Ungewissheitsurteile in der Psychophysik. *Arch. f. d. ges. Psychol.*, 1931, **80**, 273–290.
44. FERNBERGER, S. W., On Absolute and Relative Judgments in Lifted-Weight Experiments. *Amer. J. Psychol.*, 1931, **43**, 560–578.
45. FERNBERGER, S. W., GLASS, E., HOFFMAN, I., and WILLIG, M., Judgment Times of Different Psychophysical Categories. *J. Exper. Psychol.*, 1934, **17**, 286–293.
46. FERNBERGER, S. W., and IRWIN, F. W., Time Relations for the Different Categories of Judgment in the 'Absolute Method' in Psychophysics. *Amer. J. Psychol.*, 1932, **44**, 505–525.
47. FISCHEL, H., Transformationserscheinungen bei Gewichtshebungen. *Zsch. f. Psychol.*, 1926, **98**, 342–365.
48. FLETCHER, H., New Methods and Apparatus for Testing the Acuity of Hearing. *Laryngoscope*, 1925, **35**, No. 7.
49. FOLGMANN, E. E. E., An Experimental Study of Composer-Preferences of Four Outstanding Symphony Orchestras. *J. Exper. Psychol.*, 1933, **16**, 709–724.
50. FRIEDMANN, P., Die Raumschwelle der Haut beim Kinde. *Zsch. f. Psychol.*, 1927, **103**, 185–202.
51. FRITZ, M. F., Experimental Evidence in Support of Professor Thurstone's Criticism of the Phi-Gamma Hypothesis. *J. Gener. Psychol.*, 1930, **4**, 346–352.
52. FRY, F. D., HAUPT, D. D. M., and WARTENA, L., The D.L. for Cutaneous Two-Point Stimulation by the Method of Single Stimuli. *J. Exper. Psychol.*, 1933, **16**, 743–744.
53. GAHAGAN. L., On the Absolute Judgment of Lifted Weights. *J. Exper. Psychol.*, 1929, **12**, 490–501.
54. GELDARD, F. A., Brightness Contrast and Heymans' Law. *J. Gener. Psychol.*, 1931, **5**, 191–206.
55. GENGERELLI, J. A., Some Quantitative Experiments with Eidetic Imagery. *Amer. J. Psychol.*, 1930, **42**, 399–404.
56. GRAHAM, C. H., and GOLDMAN, N., Intensity and Number of Cones in Foveal Stimulation. *Amer. J. Psychol.*, 1932, **44**, 275–288.

57. GUILFORD, J. P., 'Fluctuations of Attention' with Weak Visual Stimuli. *Amer. J. Psychol.*, 1927, **38**, 534–583.
58. GUILFORD, J. P., The Method of Paired Comparisons as a Psychometric Method. *Psychol. Rev.*, 1928, **35**, 494–506.
59. GUILFORD, J. P., Some Empirical Tests of the Method of Paired Comparisons. *J. Gener. Psychol.*, 1931, **5**, 64–77.
60. GUILFORD, J. P., Racial Preferences of a Thousand American University Students. *J. Soc. Psychol.*, 1931, **2**, 179–204.
61. GUILFORD, J. P., A Generalized Psychophysical Law. *Psychol. Rev.*, 1932, **39**, 73–85.
62. GUILFORD, J. P., and DALLENBACH, K. M., The Determination of Memory Span by the Method of Constant Stimuli. *Amer. J. Psychol.*, 1925, **36**, 621–628.
63. GUILFORD, J. P., and PARK, D. G., The Effect of Interpolated Weights upon Comparative Judgments. *Amer. J. Psychol.*, 1931, **43**, 589–599.
64. HALLETT, W. N., A Psychophysical Study of Visual Gestalten. *Amer. J. Psychol.*, 1933, **45**, 691–700.
65. HAMILTON, H. C., The Effect of Incentives on Accuracy of Discrimination Measured on the Galton Bar. *Arch. of Psychol.*, 1929, No. 103. Pp. 73.
66. HECHT, S., The Relation Between Visual Acuity and Illumination. *J. Gen. Physiol.*, 1928, **11**, 255–281.
67. HECHT, S., Vision: II. The Nature of the Photoreceptor Process. In *The Foundations of Experimental Psychology*. Worcester, Mass.: Clark Univ. Press, 1929. Pp. 216–272.
68. HEVNER, K., An Empirical Study of Three Psychophysical Methods. *J. Gener. Psychol.*, 1930, **4**, 191–212.
69. HOAGLAND, H., The Weber-Fechner Law and the All-or-None Theory. *J. Gener. Psychol.*, 1930, **3**, 351–373.
70. HOUSTOUN, R. A., and SHEARER, J. F., Fechner's Law. *Nature*, 1930, **125**, 891–892.
71. HULIN, W. S., and FRAZIER, L. R., Two Comparison Stimuli with One Initial Standard in Lifted-Weights. *Amer. J. Psychol.*, 1928, **40**, 626–627.
72. HUMES, J. F., The Effect of Practice upon the Upper Limen for Tonal Discrimination. *Amer. J. Psychol.*, 1930, **42**, 1–16.
73. IRWIN, F. W., Thresholds for the Perception of Difference in Facial Expression and Its Elements. *Amer. J. Psychol.*, 1932, **44**, 1–17.
74. IRWIN, F. W., und FERNBERGER, S. W., Einige zeitliche Bestimmungen der Urteile in der Methode der einzelnen Reize. *Arch. f. d. ges. Psychol.*, 1933, **89**, 541–548.
75. JACOBS, M. H., Ueber den Einfluss des phänomenalen Abstandes auf die Unterschiedsschwelle für Helligkeiten. *Psychol. Forsch.*, 1933, **18**, 98–142.
76. JOHNSON, H. M., Some Fallacies Underlying the Use of Psychological 'Tests'. *Psychol. Rev.*, 1928, **35**, 328–337.
77. JOHNSON, H. M., Did Fechner Measure '*Introspectional*' Sensations? *Psychol. Rev.*, 1929, **36**, 257–284.
78. JOHNSON, H. M., Some Properties of Fechner's 'Intensity of Sensation'. *Psychol. Rev.*, 1930, **37**, 113–123.

79. KATZ, D., und NOLDT, F., Ueber die kleinsten vibratorisch wahrnehmbaren Schwingungen. *Zsch. f. Psychol.*, 1926, **99**, 104–109.
80. KELLOGG, W. N., Measuring Auditory Intensive Thresholds in Electrical Units. *J. Exper. Psychol.*, 1929, **12**, 240–248.
81. KELLOGG, W. N., The Graphic Method in Psychophysics. *Amer. J. Psychol.*, 1929, **41**, 456–459.
82. KELLOGG, W. N., An Improved Photometer for Measuring Brightness Thresholds in Hundredths of a Meter Candle. *J. Gener. Psychol.*, 1929, **2**, 369–377.
83. KELLOGG, W. N., An Experimental Comparison of Psychophysical Methods. *Arch. of Psychol.*, 1929, No. 106. Pp. 86.
84. KELLOGG, W. N., An Experimental Evaluation of Equality Judgments in Psychophysics. *Arch. of Psychol.*, 1930, No. 112. Pp. 78.
85. KELLOGG, W. N., The Time of Judgment in Psychometric Measures. *Amer. J. Psychol.*, 1931, **43**, 65–86.
86. KENNETH, J. H., and THOULESS, R. H., Relationship Between the Absolute and Differential Thresholds for an Auditory Stimulus. *Amer. J. Psychol.*, 1930, **42**, 389–398.
87. KIESOW, F., Ueber die Vergleichung linearer Strecken und ihre Beziehung zum Weberschen Gesetze. (Erste Mitteilung.) *Arch. f. d. ges. Psychol.*, 1925, **52**, 61–90.
88. KIESOW, F., Ueber die Vergleichung linearer Strecken und ihre Beziehung zum Weberschen Gesetze. (Zweite Mitteilung.) *Arch. f. d. ges. Psychol.*, 1925, **53**, 433–446.
89. KIESOW, F., Ueber die Vergleichung linearer Strecken und ihre Beziehung zum Weberschen Gesetze. (Dritte Mitteilung: Die Unterschiedsempfindlichkeit.) *Arch. f. d. ges. Psychol.*, 1926, **56**, 421–451.
90. KINGSBURY, B. A., A Direct Comparison of the Loudness of Pure Tones. *Phys. Rev.*, 1927, **29**, 588–600.
91. KÖHLER, W., Zur Theorie des Suksessivvergleichs und der Zeitfehler. *Psychol. Forsch.*, 1923, **4**, 115–175.
92. KÖHLER, W., Zur Psychophysik des Vergleichs und des Raumes. *Psychol. Forsch.*, 1933, **18**, 343–360.
93. KREEZER, G., The Conditions Underlying the Time and Space Errors. *Psychol. Bull.*, 1933, **30**, 685–686.
94. KRÜGER, H., Ueber die Unterschiedsempfindlichkeit für Beleuchtungseindrücke. *Zsch. f. Psychol.*, 1925, **96**, 58–67.
95. LAUENSTEIN, O., Ansatz zu einer physiologischen Theorie des Vergleichs und der Zeitfehler. *Psychol. Forsch.*, 1932, **17**, 130–177.
96. LIENAU, C. C., Singer's Definition and the Generalized Law of the Geometric Mean in Numerical Estimation. *J. Exper. Psychol.*, 1934, **17**, 189–222.
97. LINDER, F. E., A Statistical Comparison of Psychophysical Methods. *Psychol. Monog.*, 1933, **44**, No. 3, 1–20.
98. LUFKIN, H. M., The Accuracy of the Method of Constant Stimuli. *Amer. J. Psychol.*, 1927, **38**, 666–667.
99. LUFKIN, H. M., The Best Fitting Frequency Function for Urban's Lifted-Weight Results. *Amer. J. Psychol.*, 1928, **40**, 75–82.

100. MAXFIELD, F. N., The Observer: An Instrument of Precision. *Ohio J. Science*, 1927, **27**, 205–218.
101. MAXWELL, R. W., The Quantitative Estimation of the Sensation of Color. *Brit. J. Psychol.*, 1929, **20**, 181–189.
102. MCGINNIS, J. M., Electrical Control with the Method of Constant Stimuli. *Amer. J. Psychol.*, 1927, **38**, 295.
103. MOHR, F., *Psychophysische Behandlungsmethoden.* Leipzig: Hirzel, 1925. Pp. xii+493.
104. MYERS, C. S., Intensity Differences of Sensations. Proc. VIII Internat. Congr. Psychol., Groningen: P. Noordhoff, 1927, 79–84.
105. NEEDHAM, J. G., The Time-Error in Comparison Judgments. *Psychol. Bull.*, 1934, **31**, 229–243.
106. NEWHALL, S. M., Linear Interpolation *vs.* the Constant Process. *Amer. J. Psychol.*, 1927, **38**, 390–402.
107. NEWHALL, S. M., An Interpolation Procedure for Calculating Thresholds. *Psychol. Rev.*, 1928, **35**, 46–66.
108. NEWHALL, S. M., Influence of the Reporting Response on the Subject's Report. *Psychol. Bull.*, 1933, **30**, 685.
109. ONOSHIMA, U., Ueber die abhängigkeit akustischer Intensitätsschritte von einem umfassenden Tonverband. *Psychol. Forsch.*, 1928, **11**, 267–289.
110. PAULI, R., und WENZL, A., Experimentelle und theoretische Untersuchungen zum Weber-Fechnerschen Gesetz. *Arch. f. d. ges. Psychol.*, 1925, **51**, 399–494.
111. PEDDIE, W., The General Applicability of Fechner's Law in Colour Sensation. *Nature*, 1929, **124**, 791–792.
112. PETERSON, R. C., and THURSTONE, L. L., The Effect of a Motion Picture Film on Children's Attitudes Toward Germans. *J. Educ. Psychol.*, 1932, **23**, 241–246.
113. PETRAN, L. A., Pitch Discrimination Near the Upper and Lower Thresholds of Audition. *Amer. J. Psychol.*, 1933, **45**, 248–262.
114. PIÉRON, H., Des lois régissant la variation de l'intensité sensorielle en fonction de l'intensité du stimulus. *Rev. Phil.*, 1928, **106**, 261–279.
115. PONZO, M., und ANGYAL, A., Zur Systematik der Gewichtsempfindungen. *Arch. f. d. ges. Psychol.*, 1933, **88**, 629–634.
116. PRATT, C. C., The Time-Error in Psychophysical Judgments. *Amer. J. Psychol.*, 1933, **45**, 292–297.
117. PRATT, C. C., Time-Errors in the Method of Single Stimuli. *J. Exper. Psychol.*, 1933, **16**, 798–814.
118. RAZRAN, G. H. S., Conditioned Responses in Animals Other Than Dogs. *Psychol. Bull.*, 1933, **30**, 261–324.
119. RICE, S. A., Statistical Studies of Social Attitudes and Public Opinion. In *Statistics in Social Studies.* Ed. by S. A. Rice. Philadelphia: Univ. of Pennsylvania Press, 1930. Pp. 171–193.
120. RICH, G. J., Psychophysical Measurement Methods. *Psychol. Bull.*, 1925, **22**, 613–648.
121. RICH, G. J., The Measure of Approximation of Data to the Phi-Gamma Hypothesis. *Amer. J. Psychol.*, 1925, **36**, 615–620.
122. RICH, G. J., Black and Grey in Visual Theory. *Amer. J. Psychol.*, 1926, **37**, 123–128.

123. RICHARDSON, L. F., Thresholds when Sensation Is Regarded as Quantitative. *Brit. J. Psychol.*, 1928, **19**, 158–166.
124. RICHARDSON, L. F., Quantitative Mental Estimates of Light and Color. *Brit. J. Psychol.*, 1929, **20**, 27–37.
125. RICHARDSON, L. F., Imagery, Conation and Cerebral Conductance. *J. Gener. Psychol.*, 1929, **2**, 324–352.
126. RICHARDSON, L. F., A Quantitative View of Pain. *Brit. J. Psychol.*, 1933, **23**, 401–403.
127. RICHARDSON, L. F., and Ross, J. S., Loudness and Telephone Current. *J. Gener. Psychol.*, 1930, **3**, 288–306.
128. RUDISILL, E. S., Constancy of Attitude in Weight Perception. *Amer. J. Psychol.*, 1925, **36**, 562–587.
129. SHAAD, D. J., and HELSON, H., Group Presentation in the Method of Constant Stimuli as a Time-Saving Device. *Amer. J. Psychol.*, 1931, **43**, 422–433.
130. SPENCER, L. T., The Validity of Heymans' Law. *Amer. J. Psychol.*, 1925, **36**, 427–433.
131. SPENCER, L. T., The Concept of the Threshold and Heymans' Law of Inhibition. I. Correlation Between the Visual Threshold and Heymans' Coefficient of Inhibition in Binocular Vision. *J. Exper. Psychol.*, 1928, **11**, 88–97.
132. SPENCER, L. T., and COHEN, L. H., The Concept of the Threshold and Heymans' Law of Inhibition. II. Correlation of the Visual Threshold and Heymans' Coefficient of Inhibition in a Single Individual, with Uniocular Vision. *J. Exper. Psychol.*, 1928, **11**, 194–201.
133. SPENCER, L. T., and COHEN, L. H., The Concept of the Threshold and Heymans' Law of Inhibition. III. The Relation of the Threshold to Estimates of Daily Variation in 'Freshness'. *J. Exper. Psychol.*, 1928, **11**, 281–292.
134. STROOP, J. R., Is the Judgment of the Group Better than that of the Average Member of the Group? *J. Exper. Psychol.*, 1932, **15**, 550–562.
135. THOMSON, G. H., Fitting of Frequency Functions to Urban's Lifted-Weight Results. *Amer. J. Psychol.*, 1929, **41**, 70–82.
136. THURSTONE, L. L., Psychophysical Analysis. *Amer. J. Psychol.*, 1927, **38**, 368–389.
137. THURSTONE, L. L., A Law of Comparative Judgment. *Psychol. Rev.*, 1927, **34**, 273–286.
138. THURSTONE, L. L., The Method of Paired Comparison for Social Values. *J. Abn. & Soc. Psychol.*, 1927, **4**, 384–400.
139. THURSTONE, L. L., Equally Often Noticed Differences. *J. Educ. Psychol.*, 1927, **18**, 289–293.
140. THURSTONE, L. L., A Mental Unit of Measurement. *Psychol. Rev.*, 1927, **34**, 415–423.
141. THURSTONE, L. L., Three Psychophysical Laws. *Psychol. Rev.*, 1927, **34**, 424–432.
142. THURSTONE, L. L., Attitudes Can Be Measured. *Amer. J. Soc.*, 1928, **33**, 529–554.
143. THURSTONE, L. L., An Experimental Study of Nationality Preferences. *J. Gener. Psychol.*, 1928, **1**, 405–425.

144. THURSTONE, L. L., The Measurement of Opinion. *J. Abn. & Soc. Psychol.*, 1928, **22**, 415–430.
145. THURSTONE, L. L., The Phi-Gamma Hypothesis. *J. Exper. Psychol.*, 1928, **11**, 293–305.
146. THURSTONE, L. L., Fechner's Law and the Method of Equal-Appearing Intervals. *J. Exper. Psychol.*, 1929, **12**, 214–224.
147. THURSTONE, L. L., Theory of Attitude Measurement. *Psychol. Rev.*, 1929, **36**, 222–241.
148. THURSTONE, L. L., Commentary. In *Statistics in Social Studies.* Ed. by S. A. Rice. Philadelphia: Univ. of Pennsylvania Press, 1930. Pp. 192–196.
149. THURSTONE, L. L., Rank Order as a Psychophysical Method. *J. Exper. Psychol.*, 1931, **14**, 187–201.
150. THURSTONE, L. L., The Indifference Function. *J. Soc. Psychol.*, 1931, **2**, 139–167.
151. THURSTONE, L. L., The Measurement of Social Attitudes. *J. Abn. & Soc. Psychol.*, 1931, **26**, 249–269.
152. THURSTONE, L. L., Stimulus Dispersions in the Method of Constant Stimuli. *J. Exper. Psychol.*, 1932, **15**, 284–297.
153. THURSTONE, L. L., and CHAVE, E. J., *The Measurement of Attitude.* Chicago: Univ. of Chicago Press, 1929. Pp. x+97.
154. TIFFIN, J., Some Aspects of the Psychophysics of the Vibrato. *Psychol. Monog.*, 1931, **41**, No. 4, pp. 153–200.
155. TINKER, M. A., The Statistical Limen versus the Average as a Measure of Visual Apprehension. *J. Exper. Psychol.*, 1934, **17**, 105–112.
156. TINKER, M. A., Methods Employed for Ranking Stimuli. *J. Gener. Psychol.*, 1931, **5**, 53–63.
157. TRAVIS, R. C., The Diagnosis of Character Types by Visual and Auditory Thresholds. *Psychol. Monog.*, 1926, **36**, No. 2, pp. 18–37.
158. TRUMAN, S. R., and WEVER, E. G., The Judgment of Pitch as a Function of the Series. *Univ. Calif. Publ. Psychol.*, 1928, **3**, 215–223.
159. TURNER, W. D., Intra-Serial Effects with Lifted Weights. *Amer. J. Psychol.*, 1931, **43**, 1–25.
160. URBAN, F. M., The Accuracy of the Method of Constant Stimuli. *Amer. J. Psychol.*, 1927, **38**, 252–256.
161. URBAN, F. M., Die Methode des durchschnittlichen Fehlers. *Arch. f. d. ges. Psychol.*, 1930, **74**, 141–162.
162. URBAN, F. M., Die Untersuchung von W. N. Kellogg über die akustische und optische Unterschieds-Schwelle. *Arch. f. d. ges. Psychol.*, 1930, **74**, 275–282.
163. URBAN, F. M., The Future of Psychophysics. *Psychol. Rev.*, 1930, **37**, 93–106.
164. URBAN, F. M., Die Verallgemeinerung der Konstanzmethode. *Arch. f. d. ges. Psychol.*, 1931, **80**, 167–178.
165. URBAN, F. M., Ueber die Methode der gleichen Abstufungen. *Arch. f. d. ges. Psychol.*, 1931, **80**, 291–312.
166. URBAN, F. M., Die Messbarkeit psychischer Mannigfaltigkeiten. *Arch. f. d. ges. Psychol.*, 1932, **85**, 229–236.

167. URBAN, F. M., The Weber-Fechner Law and Mental Measurement. *J. Exper. Psychol.*, 1933, **16**, 221–238.
168. VOLKMANN, J., The Method of Single Stimuli. *Amer. J. Psychol.*, 1932, **44**, 808–809.
169. WADA, Y., Der Zeitfehler beim Suksessivvergleich der Tonhöhen. *Jap. J. Psychol.*, 1932, **7**, 505–537.
170. WALTER, P., Ueber den Einfluss des Spannungsablaufes im Muskel auf die Gewichtswahrnehmung. *Zsch. f. Psychol.*, 1927, **104**, 97–108.
171. WALTON, W. E., GUILFORD, R. B., and GUILFORD, J. P., Color Preferences of 1,279 University Students. *Amer. J. Psychol.*, 1933, **45**, 322–328.
172. WENDT, G. R., Two Industrial Applications of a Psychophysical Method. *J. Appl. Psychol.*, 1932, **16**, 269–276.
173. WERNER, H., Ueber die Intensität der Empfindungen. Proc. VIII Internat. Congr. Psychol. Groningen: P. Noordhoff, 1927, 85–93.
174. WERTHEIMER, M., Zum Problem der Schwelle. Proc. VIII Internat. Congr. Psychol. Groningen: P. Noordhoff, 1927, 447.
175. WEVER, E. G., and TRUMAN, S. R., The Course of the Auditory Threshold in the Presence of a Tonal Background. *J. Exper. Psychol.*, 1928, **11**, 98–112.
176. WEVER, E. G., and ZENER, K. E., The Method of Absolute Judgment in Psychophysics. *Psychol. Rev.*, 1928, **35**, 466–493.
177. WHITE, R. K., and LANDIS, C., Perception of Silhouettes. *Amer. J. Psychol.*, 1930, **42**, 431–435.
178. WOODROW, H., Weight-Discrimination with a Varying Standard. *Amer. J. Psychol.*, 1933, **45**, 391–416.
179. WOODWORTH, R. S., Psychological Data Pertaining to Errors of Observation. *International Critical Tables*, 1926, **1**, 92–95.
180. ZINNER, E., Die Reizempfindungskurve. *Zsch. f. Sinnesphysiol.*, 1930, **61**, 247–266.
181. ZINNER, E., Ueber die Darstellung der Reizempfindungskurve. *Zsch. f. Sinnesphysiol.*, 1933, **64**, 175–176.
182. ZÖLLER, W., *Formeln und Tabellen zur Errechnung des mittleren Fehlers.* Berlin, 1925. Pp. 32.

KELLOGG, W. N., and KELLOGG, L. A. *The Ape and the Child.* New York: McGraw-Hill Book Company, 1933. Pp. xiv+341.

In this book Professor and Mrs. Kellogg have summarized the results of a unique psychological experiment—that of taking a young ape into a human environment for a period of nine months during which time its care and training were in all respects identical with those of the human infant belonging to the household. The subject of the experiment, a young female chimpanzee, was born in captivity in the Abreu Colony in Cuba. She was removed from the ape colony at the age of 7½ months and was taken to the Kelloggs' home where she was gradually initiated into the human mode of living. Here she had as her constant companion and playmate the author's infant son, 2½ months older than the ape. The two babies were treated in as nearly identical a fashion as possible. They were dressed alike, given similar toys, rode in the same perambulator, were talked to, played with and caressed in the same manner and in equal degree. Every attempt was made to keep the psychological as well as the material environment of the ape the same as that of the child. In this respect the experiment probably stands alone. Apes have been kept as pets before and, as every vaudeville attendant can testify, they have been found capable of learning many clever tricks. At least one chimpanzee has been reared from birth to the age of nine months by human beings [1] but although reared by humans she was not reared as a human. The Jacobsens' study, therefore, does not parallel the Kellogg experiment, although certain interesting and valuable comparisons can be made between them.

The major part of the present report is devoted to a topical comparison of the development and behavior of the child and the ape. Physical growth, reflexes, dentition and the like are described rather briefly, since the main purpose of the study was psychological rather than anthropometric. Routine habits of eating and sleeping, motor and sensory development, social and emotional behavior, memory and language development are for the most part handled through

[1] Jacobsen, C. F., and Jacobsen, M. M., Development of an Infant Chimpanzee During Her First Year. *Comp. Psychol. Monog.*, 1932, **9**, Serial No. 41, Pp. 94.

descriptions of everyday behavior. Careful accounts are given of a number of simple learning experiments in which the conditions were made identical for the two subjects, and there is a detailed analysis of month by month performances on the Gesell tests.

Interpretation of the findings is difficult. Although the child was 2½ months older than the ape, the latter was physiologically much in advance of the child. As the authors are careful to point out, many of the observed differences in the behavior of the two infants may be accounted for on the basis of differences in bodily strength and in anatomical structure. For example, the ape was decidedly inferior to the child in fine prehension and in piling blocks—tasks for which the form of the ape's hand is not well adapted. The ape, however, was much superior to the child in climbing, jumping, and in all tasks dependent chiefly upon strength and agility. Surprisingly enough, the ape also excelled in most of the experiments on learning and problem solving. This is not so easily explained in terms of structural differences but may, as the authors point out, be attributable at least in part to differences in motivation, for the ape showed far more interest in securing the bits of fruit or cookie which were used as incentives than did the child. For a time the ape excelled in comprehension of human speech, but at all ages the child's repertoire of vocal sounds was far more extensive and varied than that of the ape.

At the end of nine months the experiment was brought to a close and the ape was returned to the anthropoid colony. To the reviewer it seems very regrettable that the study was not carried further. It is to be expected that with increasing age the initial advantages resulting from the more rapid physiological development of the chimpanzee would gradually be lost and the child would forge ahead, but while there was some slight indication during the latter part of the experiment that such a process was taking place, the evidence is far from being clear. On the Gesell tests, for example, which were given to both infants at monthly intervals, a somewhat questionable method of scoring based upon cumulative gains suggested that the child was improving at a more rapid rate than the ape, but at the termination of the experiment the ape still scored appreciably in advance of the child's performance at a corresponding age. Had the same test results been obtained for two human infants I question whether any competent psychologist would be willing to express a confident opinion that the one with the lower score, who, however, showed some tendency toward a faster rate of gain, would eventually prove to be superior to the other. The conclusion in the present instance

seems to have been determined more by the knowledge that the ape was an ape than by the evidence actually obtained.

Whether or not the ape would have learned human speech if she had been retained longer in the human environment is another question of vital significance which remains unanswered. True, at the age of 16½ months she had made little or no observable progress in verbal expression and her repertoire of vocal sounds appeared to be little, if at all, more extensive than that reported for other young chimpanzees. But the human infant who was 2½ months older had likewise made little progress in actual speech at the time the experiment was brought to a close. The absence in the ape of the expressive jargon so characteristic of the vocalizations of human infants during the latter part of the prelinguistic period renders it somewhat less probable that true speech would have developed later but does not afford conclusive evidence that it might not have appeared.

All in all, one lays the book down with a feeling of disappointment that an experiment so well begun should have been discontinued before crucial evidence as to the modifiability of the ape's behavior had been secured. It is to be hoped that at some time in the future another study of this kind may be undertaken and carried through to completion.

FLORENCE L. GOODENOUGH.

University of Minnesota.

KATZ, DAVID. Zur Psychologie des Amputierten und seiner Prothese.[1] *Beihefte zur Zeitschrift für angewandte Psychologie,* 1921, **25**. Pp. vi+118.

This monograph embodies the results of psychological studies of soldiers with amputated limbs. The studies were made at the Psychological Research Station for Artificial Limbs in Hanover and at the Psychological Institute of the University of Rostock. They were undertaken by Dr. Katz under government appointment in April, 1918, when the great number of amputations from war injuries made rehabilitation an increasingly great social and economic problem.

The point of departure in the study was the fact that artificial limbs usually proved to be more of a hindrance to the wearer than a practical help. As nearly as could be ascertained, only about 13 per cent of those with artificial limbs wore them when working.

[1] Although this monograph appeared in 1921, it is felt that American readers will welcome this review because of the unusual character of the investigations presented. THE EDITOR.

The difficulty seemed to be largely that of a loss of cutaneous and muscular sensitivity through the heavy padding and the leather casing over the stump of the amputated limb. The object of the study was an investigation of the stump in the hope that a psychological basis might be found for a more satisfactory fitting of artificial limbs or other appliances.

The subjects for the study were 102 soldiers, of which only 3 had leg amputations. The rest included all varieties of arm amputations. The method employed, in general, was that of a free (introspective) description by each subject of the sensations of the stump, including illusions relating to the missing limb, followed by a detailed experimental (objective and introspective) study of the sensory and motor ability of the stump.

All subjects reported that they had or had had illusions relating to the existence of the missing part of the limb. These illusions represented in various forms a "phantom limb" which appeared more sharply and continuously in consciousness than did the perception of the remaining sound limb. The sensations on which these illusions were based seemed to be those of muscle and tendon strain. They were described as like sensations from a sleeping limb, like a tight cuff around the hand, like a gummy mass in the palm which was being squeezed through the fingers, and like the pain from pressure of an iron or wood object in the palm.

The reality and intensity of the "phantom limb" sensations were tested and confirmed by a method called the "method of equivalents". Stimuli such as bandages or clamps were applied to corresponding parts of the sound arm until the sensations seemed to be equally intense. The liveliness of the illusions was also shown by the fact that the subjects, when in an inattentive state, would reach for a falling object with the "phantom" hand. These illusions were reported as appearing immediately after the amputation by 57 subjects whose amputations followed closely after the injury. There was great variation in the duration of the illusions, but all seemed to fade gradually and slowly with time.

The structure of the "phantom limb" varied considerably with different subjects. Of 22 cases of forearm amputation, the missing part of the arm was felt in only one case; the fingers were felt by all; and the hand and wrist followed these in the frequency of their representation. The limb was usually smaller than life-size, the hand being frequently of the size of a child's hand. The position and length of the "phantom limb" were ascertained by asking the subject to indicate the position of the "phantom" hand with his remain-

ing (normal) hand. In 3 of the 22 cases questioned the "phantom limb" appeared to be of normal length. In 7 of the 22, it seemed to be telescoped wholly within the stump. Most of the cases lay between these extremes, revealing a shortened arm. The limb was frequently located at an angle with the stump.

The sensations from the stump, referred to the "phantom limb", seemed to be stimulated by muscle tonus, movement of clothing on the skin, differences in temperature, position of the bones, etc. The illusions were more clearly in consciousness when the stump was passive; they usually disappeared while the stump was moving.

The "phantom limb" illusions had many of the characteristics of visual hallucinations, and were altered by visual perceptions. When the "phantom limb" was felt to be within the stump, an actual stimulus on the stump was felt to be there alone. When the "limb" occupied a space beyond the stump, it seemed to bend or telescope to avoid an approaching solid object. The perception of the "limb" was sharper when the subject was not looking at the space it would occupy.

Movements of the "phantom limb" were reported as felt by most of the subjects. Movements of parts could be felt as follows: 44 of 53 cases reported illusions of finger movements; 19 reported wrist movements; and only 5 of 23 upper-arm cases felt movements of the elbow. The length of time ensuing after the operation seemed to bear little or no relation to the motion illusions.

The part played by the peripheral and central divisions of the nervous system in bringing about the illusions has been interpreted variously. The author holds the view that the illusions are aroused and their intensity graded by the nervous stimuli of the stump, but the form (content) of the illusion is a function of the central nervous system.

There has also been a difference of opinion in regard to the innervation of the muscles of the stump: whether the antagonistic action of the triceps and biceps is due to nervous linkage, or whether it is due primarily to their anatomic connection with the bones of the forearm. The fact of significance in this study is that when the biceps and triceps are cut loose from their connection with the forearm, they can be trained to contract at will, either separately or simultaneously. Relaxation of the muscles followed contraction as a passive process rather than through active innervation accompanying contraction of the antagonistic muscle.

The rôle of the "phantom limb" in dream-life varied greatly. Of 25 questioned, only 14 were able to give any definite information.

Of these, some reported that the missing limb seemed to be quite normal, others that it seemed to be missing from their bodies.

The second part of the investigation included an analytical study of the sensations from the stump, the degree of usefulness of the sensations from the stump in perception (praktische Leistungsfähigkeit), and a detailed study of those upper-arm stumps on which had been performed the Sauerbruch operation.[2]

The pressure sensations were determined with a von Frey esthesiometer for the stump and the normal arm. The average threshold was decidedly lower near the scar than on other parts of the stump or on corresponding parts of the normal arm. This difference was judged to be a matter of attention rather than a physiological difference in sensitivity.

The two-point threshold was determined with Spearman's esthesiometer in four to six places on the stump and the corresponding areas of the normal arm. The average threshold in upper-arm amputations (8 cases) was 52 mm. on the stump, and 71 mm. for corresponding areas on the other arm. With forearm amputations (11 cases), the averages were 37 mm. and 52 mm. This difference is also explained as due to central rather than peripheral factors.

Accuracy in localization of contact stimuli was tested by stimulation of rows of points marked on corresponding dorsal areas of the two arms. These were tested under various conditions and the errors measured. It was found that the localization error was in the direction of the shoulder (proximal) and was greater on the stump than on the other arm. This was partly due, perhaps, to the shortening of the arm and loss of mobility, and partly to the effect of the foreshortened "phantom limb".

The reaction-time for contact stimulation (33 subjects) was not significantly different for stimulations of the stump and the normal arm. The time was greater with arm stimulation in 18 subjects and less with 14, but the difference was small.

The practical usefulness of the sensitivity of the stump was tested by comparing the perception of familiar objects through contact with the stump and the remaining hand. Thirty-three subjects were tested for recognition of familiar surfaces which were rubbed gently

[2] The Sauerbruch operation consists of cutting canals transversely through the biceps and triceps muscles near the end of the stump. Through each canal, which is padded with soft leather, is put an ivory rod, from the ends of which an iron loop is suspended. This makes it possible to attach a weight to the muscle. After recovery from the operation, patients are put on regular exercise to strengthen the muscles and train them in use under the new conditions.

on the skin of the stump and the fingers. The materials were cotton, silk, woolen and linen cloth, glass, glossy paper, coarse wrapping paper, sand paper, etc. Of 12 surfaces presented, an average of 9 were recognized correctly through hand stimulation and 6 through stump stimulation. There was no difference between forearm and upper-arm sensitivity. Discrimination improved markedly with practice, becoming perfect in some subjects with the materials used.

The perception of numbers was similarly tested with nails driven into a board in patterns. The spaces between nails were about 3 cm. Perception was poor with stump stimulation. Of 36 cases tested, 17 could not perceive 2 nails, 12 distinguished 2, 7 distinguished 3 or 4, and none could distinguish as many as 5.[3]

Perception through the stump of such familiar objects as a hammer, pencil, book, small cardboard box, etc., was so poor that no statistical evidence was included in the monograph.

The triceps and biceps muscles on which had been performed the Sauerbruch operation were tested separately for stretch, lifting power and weight discrimination. It was found in the weight discrimination tests that the threshold of discrimination was higher for the biceps than for the hand, but that the error in judgment tended to be in the same direction. The weight increment ratio decreased as the standard weight was increased. The standard weights used were 225, 550, and 1,100 gms.

It was also determined that, although the lifting of weights by the hand is usually accompanied by relaxation of the triceps, the discrimination of weights lifted by the triceps alone was not materially different from that with the biceps. The threshold for weights lifted with the shoulder was higher than that found with the muscles.

The threshold of discrimination for the passive muscle was found to be lower in each case than for the same active muscle. Apparently, the pressure and strain sensation thresholds were lower when the muscle was passive.

The pressure sensations in the muscle canals were tested with rods of various diameters. Since the smaller rods were supported by less surface, they exerted more pressure for the same weight than the larger rods. The weights were correspondingly estimated as heavier, but the difference was not strictly proportional to the difference in area in contact with the rod.

A test of the impressions of magnitude of weights lifted with the

[3] The failure to obtain results in this experiment was possibly due to the fact that the distance between nails was appreciably below the two-point threshold.

Sauerbruch muscles in comparison with the weights lifted with the remaining forearm was finally made. The arm was bent at an angle of 90 degrees so that the pull should be largely on the biceps. The weight seemed to be lighter in each case as lifted directly by the muscle. This difference seemed to be mainly a matter of the direct pull of the muscle as compared with the pull at the end of a lever arm. The weight comparisons varied closely with the length of the lever arm as was shown by moving the weight along the forearm.

This monograph shows the results of an exhaustive study made with exceptional care. Caution should be observed, however, in drawing general conclusions from some of the experimental results, because of the small number of subjects available. The material is of special interest because it concerns a field about which very little has been written in English.

EVELYN F. FRENCH.

University of Oregon.

LUND, FREDERICK H. *Psychology.* New York: The Ronald Press Company, 1933.

The numerous textbooks on elementary general psychology that have appeared in the past few years are similar in one respect—they represent so many attempts to make the subject matter of psychology comprehensible, interesting and profitable for the beginning student. Some seem to have achieved this in a limited way, and it is true that one may find useful original contributions in any book. But the fact that so many books are appearing which do not differ in fundamental respects from the older ones may be taken to indicate that the older treatments are generally satisfactory or that something is lacking, the precise nature of which is not clearly understood. There is of course a formal aspect in the treatment provided by any book, and there are many details in theory and fact that need revision from time to time, but the major task is that of presenting an elementary system of psychology, even though crudely done, that will begin to explain vital phenomena as they actually occur. Man alive and in action is the concern of the student. S—R diagrams and a host of other diagrams of similar nature, recalling as they do mathematical symbols to the non-mathematical and even hostile mind, are mainly devices for a teacher. The student is interested in man as the *psychologist* sees him.

Professor Lund's point of view is empirical and biological. The study of psychology should be prosecuted by the same methods as are employed in the physical and biological sciences. The general

way in which this point of view is developed is clearly shown on page 30 where the elements involved in a simple adjustment, the S—R unit, are enumerated. These include stimulation, excitation, afferent transmission, coördination, efferent transmission, and response. The problem of adjustment, with this scheme of analysis serving as a point of departure, is discussed under the following headings: Characteristics of Human Adjustment, The Receptors, The Effectors, The Central Processes, Social and Verbal Behavior, Perceptual Responses, Thinking, Sleeping, Dreaming and Imagining, Individual Differences, and Intelligence.

The psychological and logical difficulties one encounters in following the biological methods too closely are sufficiently well known to deserve no comment. One can hardly object to explaining man from a purely biological point of view, so far as this can be done. The trouble with this sort of system is that it does not explain man from the complex psychological point of view. Professor Lund has, however, done a good job of applying the method. The book is clearly written and very well illustrated. Another feature of the book is to be found in the numerous references at the end of each chapter.

JAMES VAUGHN.

University of Cincinnati.

LOUTTIT, C. M. *Handbook of Psychological Literature.* Bloomington, Indiana: The Principia Press, 1932.

No graduate student knows or even needs to know everything this manual has to offer about the printed sources of information in psychology but every graduate student, one predicts, would profit by its perusal. There are valuable hints on notetaking and the preparation of bibliographies, a study of journals, of general reference works, and of the usually neglected institutional publications, a particularly useful chapter on "Literature Guides", and a brief account of library principles. Little attention is given to searching books, though this is a difficult and little appreciated art.

The appendix lists 1,083 journals in psychology and related subjects, cross-indexed under some sixty topics. These topics reflect the fields to which the journals are more or less ostensibly devoted and are therefore of decidedly uneven importance. If the limitations of the index are recognized, however, it should prove a useful starting point.

HORACE B. ENGLISH.

The Ohio State University.

LILLIE, RALPH S. *Protoplasmic Action and Nervous Action.* Second Edition. Chicago: University of Chicago Press, 1932. Pp. xiii+417.

The first edition of this book, which was issued in 1923, has been generally recognized as an outstanding contribution to one of the most fundamental of biological problems. The phenomena of stimulation and transmission of excitation in protoplasm and nerve have been examined in reference to the physico-chemical organization of protoplasm. Professor Lillie has drawn upon an extensive experimental literature in addition to his own highly significant researches.

Although the publishers announce that this is a "revised edition" in which the author "brings his materials and bibliography up to date", there is surprisingly little change from the first edition. The text remains almost unaltered, while at the end of each chapter is appended a footnote which cites a few papers designed to serve as an introduction to the experimental literature since 1923.

DONALD G. MARQUIS.

Yale University.

BENTLEY, MADISON. *The New Field of Psychology.* New York: D. Appleton-Century Company, 1933. Pp. xvi+439.

Although this volume is offered as an elementary textbook it constitutes a publication of capital importance. Written by the Sage Professor of Psychology at Cornell University, it symbolizes the tremendous changes in the psychological atmosphere at that institution since Titchener's Textbook was in vogue.

What then are these tremendous changes? No less than the dropping of 'mind', 'consciousness', and 'introspection' from the vocabulary of the psychologist.

We may therefore take this volume to be one of the outstanding milestones in the progressive movement of psychology toward objectivity. The rapidity with which psychology is changing we may gather from a comparison of the present volume with the same author's *The Field of Psychology* published in 1924, of which the present work is a revision and simplification. Bentley says that in the older book he had already insisted upon a functional unity in man, but still could not avoid the intermixture of magical and natural causes or the monstrous mind-body problem.

In this book the author believes himself rid of all these encumbrances. So radically changed has been the envisagement of the subject that the appeal to mental processes, states, and forces has

been excluded. Psychology becomes for him the study of bodily functions of which there are two kinds: one including digestion, circulation, etc., the other, perceiving, comprehending, remembering, etc.

Has Bentley then turned behaviorist? Not at all. He is still opposed to behaviorism because it ignores the partial independence of the organic system, which has its own powers of initiation, direction, and control. Moreover, Bentley believes that there is a definite hiatus between the behavioristic study of physiological activities and such phenomena as perceiving, understanding, remembering, etc.

Having rejected mentalism and behaviorism does the author turn to organismic psychology? As we shall see, the answer is both no and yes.

In the forefront of Bentley's exposition stands the concept of function and product. Psychological functions, like physiological functions, he says in the preface, have issues. Just as digesting turns out nutritive and waste substances, so languages, histories, arts, beliefs, etc., are products which issue from the psychological functions. This all sounds plausible, but now turn to the actual psychological functions—for example, perceiving. What is the nature of the function and of the product? The function is the conventionally acknowledged unknown, but hopefully looked for psychological (*sic*) operation of the receptors, brain, and other organic structures. "As for perceiving, we shall emphasize those bodily means which are especially employed when we do perceive" (p. 92), although further he says, "the real key to the activities of the brain is still to be found" (p. 118).

Perceived things are the products of this unknown activity. "*We should not think of the stimulus exciting the receptor as the perceived object but always as a form of energy described in terms of the physicist.* Thus the visual stimulus is not the tree or the rainbow (which is the *product* of the perceptive function) but rather the pattern of the radiant energy . . ." (p. 94, author's italics). The catalog of these products gives the great qualitative multiplicity—that is, the visual, auditory, etc., qualities, or the patterns of apprehension.

What is this but the same mentalistic doctrine of the former book? Despite Bentley's constant reiteration that he does not deal with mental processes or sensations, he is still a mentalist. Is not the color created by the receptor or brain the old experience or psychic stuff or process? Not only does this conception clash with the author's commendable intentions to get rid of mental processes, but

with the genuinely objective (organismic) conceptions and descriptions in the book.

For example, in describing the details of perception there is another story both implied and expressed. "Compare the child's apprehension of a plain gold band with that of the bride" (p. 375). Here as in the description of ways in which perceptual behavior develops we have a statement of objective interactions of organisms with objects. There is no necessity, as Bentley shows there is no possibility, of describing observable happenings in terms of hypothetical eye and brain functions and the abstractions of physics. Only by assuming the existence of the gold band in its physical, chemical, and anthropological aspects can one describe the interactions of an organism with it in a satisfactorily objective or scientific manner.

Nor indeed does Bentley overlook this point. For in discussing the governors of conduct he says, "The very objects of apprehension and the subjects of comprehension and thinking are largely provided from this same human source" (p. 27). As this source consists of a series of extra-organic governors, must they not be objectively existing phenomena?

In sum, this volume may be regarded either as an unintegrated statement of opposed concepts or as merely verbally objective but actually mentalistic. But there is still the third possibility already suggested of looking upon this book, despite the intellectual lag of some of its material, as having in a measure attained that objectivity toward which psychology must inevitably proceed. Because it indicates the direction of that objectivity it merits the careful study of all those who are interested in understanding the fundamentals of psychology.

J. R. KANTOR.

Indiana University.